HAL
open science

Semiclassical and spectral analysis of oceanic waves

Christophe Cheverry, Isabelle Gallagher, Thierry Paul, Laure Saint-Raymond

To cite this version:

Christophe Cheverry, Isabelle Gallagher, Thierry Paul, Laure Saint-Raymond. Semiclassical and spectral analysis of oceanic waves. Duke Mathematical Journal, 2012, 161 (5), pp.845-892. hal-00481593v3

HAL Id: hal-00481593

https://hal.science/hal-00481593v3

Submitted on 14 Jun 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

SEMICLASSICAL AND SPECTRAL ANALYSIS OF OCEANIC WAVES

CHRISTOPHE CHEVERRY, ISABELLE GALLAGHER, THIERRY PAUL,
AND LAURE SAINT-RAYMOND

Abstract. In this work we prove that the shallow water flow, subject to strong wind forcing and linearized around an adequate stationary profile, develops for large times closed trajectories due to the propagation of Rossby waves, while Poincaré waves are shown to disperse. The methods used in this paper involve semi-classical analysis and dynamical systems for the study of Rossby waves, while some refined spectral analysis is required for the study of Poincaré waves, due to the large time scale involved which is of diffractive type.

1. Introduction

The problem we consider is motivated by large-scale oceanography: the main physical phenomenon leading this study is the existence of persistent oceanic eddies, which are coherent structures of vortex type, spreading over dozens of kilometers and propagating slowly over periods from one year to one decade. These structures have been observed long past by physicists [16, 17, 19, 26, 27] who gave heuristic arguments (reproduced below) to explain their formation due both to wind forcing and to convection by a macroscopic zonal current. Giving a (much less precise) mathematical counterpart of those arguments, even at a linear level, requires careful multiscale analysis and rather sophisticated tools of semiclassical and microlocal analysis. In this paper we simplify the model by considering particular macroscopic currents, which are stationary solutions of the forced equations. This allows to exhibit trapped Rossby waves, by solving the dynamics associated with an appropriate integrable Hamiltonian system. We prove also that the other waves produced by the dynamics, namely Poincaré waves, disperse on the same time scales (which turn out to be of diffractive type).

1.1. **Physical observations.** Simple observations show that **large-scale ocean dynamics** can be decomposed as the sum of the solid-body rotation together with the Earth, convection by macroscopic currents (such as the Gulf Stream in the North Atlantic, the Kuroshio in the North Pacific, Equatorial or Circumpolar currents), and motion on smaller geographical zones, due for instance to the fluctuations of the wind and more generally to the coupling with the atmosphere. While the spatial extent of macroscopic currents is of the order of a hundred to a thousand kilometers, those fluctuations are typically on dozens of kilometers. We therefore expect eddies to be particular forms of those fluctuations. The point is to understand why they are quasi-stationary, or in other words why they do not disperse as other waves. At this stage we have to describe briefly the **different kinds of waves** that can be found in the ocean as linear responses to exterior forcing. They are usually classified into two families,

2010 *Mathematics Subject Classification.* 35Q86; 76M45; 35S30; 81Q20.

Key words and phrases. Semiclassical analysis; microlocal analysis; integrable systems; Bohrn-Sommerfeld quantization; Geophysical flows.

depending on their typical period and on their dynamical structure. The exact dispersion relation of all these waves can be computed explicitly [4, 10, 13, 14, 24] in simplified cases (no convection, linear approximation of the Coriolis parameter).

- **Poincaré waves**, the period of which is of the order of a day, are fast dispersive waves. They are due to the Coriolis force, that is to the rotation of the Earth ;
- **Rossby waves** propagate much slower, since the departure from geostrophy (that is equilibrium between pressure and Coriolis force) is very small. They are actually related to the variations of the Coriolis parameter with latitude. In particular, they propagate only eastwards.

The heuristic argument leading to the existence of quasi-stationary coherent structures is then as follows (as suggested by physicists): the wind forcing produces waves, in particular Rossby waves which would propagate, in the absence of convection, with a speed comparable to the bulk velocity of the fluid $\bar{v} \sim 10\,\mathrm{ms}^{-1}$; the convection by zonal flow may then stop the propagation, creating ventilation zones which are not influenced by external dynamics, in particular by continental recirculation. We are then led to studying wave propagation under the coupled effects of the pressure, the Coriolis force and zonal convection, that is to studying a system of linear PDEs with non constant coefficients.

1.2. **The model.** The system we will consider is actually a toy model insofar as many physical phenomena are neglected. Our aim here is only to get a qualitative mechanism to explain the trapping of Rossby waves. More precisely, we consider the ocean as an **incompressible, inviscid fluid with free surface** submitted to gravitation and wind forcing, and further make the following classical assumptions : the density of the fluid is homogeneous $\rho = \rho_0 =$ constant ; the pressure law is given by the hydrostatic approximation $p = \rho_0 g z$; the motion is essentially horizontal and does not depend on the vertical coordinate, leading to the so-called **shallow water approximation**. For the sake of simplicity, we shall not discuss the effects of the interaction with the boundaries, describing neither the vertical boundary layers, known as Ekman layers, nor the lateral boundary layers, known as Munk and Stommel layers. We consider a purely horizontal model, and assume an infinite domain for the longitude (omitting the stopping conditions on the continents) as well as for the latitude (this may be heuristically justified using the exponential decay of the equatorial waves to neglect the boundary). The evolution of the water height h and velocity v is then governed by the **Saint-Venant equations with Coriolis force**

$$\begin{aligned} \partial_t(\rho_0 h) + \nabla\cdot(\rho_0 h v) &= 0 \\ \partial_t(\rho_0 h v) + \nabla\cdot(\rho_0 h v\otimes v) + \omega(\rho_0 h v)^\perp + \rho_0 g h \nabla h &= \rho_0 h \tau \end{aligned} \tag{1.1}$$

where ω denotes the vertical component of the Earth rotation vector Ω, $v^\perp := (-v_2, v_1)$, g is the gravity and τ is the - stationary - forcing responsible for the macroscopic flow. It depends in particular on time averages of the wind forcing, temperature gradients and topography. The equations are written in cartesian coordinates (x_1, x_2), where x_1 corresponds to the longitude, and x_2 to the latitude (both will be chosen in $\mathbf{R}$). The vertical component of the Earth rotation is therefore $\Omega\sin(x_2/R)$, where R is the radius of the Earth, but it is classical in the physical literature to consider the linearization of ω (known as the betaplane approximation) $\omega(x_2) = \Omega x_2/R$; most of our results will actually hold for more general functions ω, but in some situations we shall particularize the betaplane case in order to improve on the results. In

order to analyze the **influence of the macroscopic convection** on the trapping of Rossby waves, we will consider small fluctuations around the stationary solution

$$h = \bar{h}, \quad \nabla \cdot (\bar{v} \otimes \bar{v}) + \omega \bar{v}^{\perp} = \tau, \quad \operatorname{div} \bar{v} = 0,$$

where $\bar{h}$ is a constant. Physical observations show that the nonlinear convection term is essentially negligible compared to the Coriolis term, so that the previous equation is nothing else than the Sverdrup relation (see [27]).

1.3. **Orders of magnitude and scaling.** Let us introduce the observation length, time and velocity scales l_0 (of the size of the radius of the Earth R), t_0 and v_0, and the nondimensional variables $\tilde{x} = x/l_0$, $\tilde{t} = t/t_0$, and $u = (v - \bar{v})/v_0$. We also define the typical height variation δh and the corresponding dimensionless variable $\eta = (h - \bar{h})/\delta h$. We denote by v_c the typical value of the velocity of the macroscopic current: $\bar{u} = (\bar{u}_1, 0) = \bar{v}/v_c$. The length scale l_0 and the convection velocity v_c are fixed by the macroscopic flow: typical values for the Gulf Stream are $l_0 \sim 10^4\,\mathrm{km}$ and $v_c \sim 10\,\mathrm{ms}^{-1}$. As we are interested in structures persisting during many months, a relevant choice for the observation time scale is $t_0 = 10^6\,\mathrm{s}$ ($\sim 0,38\,\mathrm{months}$). The associated **Rossby number** is then $\mathrm{Ro} := 1/(t_0|\Omega|) = 0.01$, recalling that $|\Omega| = 7.3 \times 10^{-5}\mathrm{s}^{-1}$. The variations of water height which can be observed are typically of the order $\delta h \sim 1\,\mathrm{m}$ to be compared to $\bar{h} \sim 10^3\,\mathrm{m}$. The influence of gravity (through hydrostatic pressure) is measured by the **Froude number** $\mathrm{Fr}^2 := (v_0 l_0)/(t_0 g \delta h) \sim 0.1$, considering namely fluctuations of order $v_0 \sim 0.1\mathrm{ms}^{-1}$. Defining $\varepsilon := \mathrm{Fr}^2$ and dropping the tildas (note that as often in Physics, ε is not really a very small number), we therefore end up with the following scaled system

$$\begin{aligned} \partial_t \eta + \frac{1}{\varepsilon} \nabla \cdot u + \bar{u} \cdot \nabla \eta + \varepsilon^2 \nabla \cdot (\eta u) = 0\,, \\ \partial_t u + \frac{1}{\varepsilon^2} b u^{\perp} + \frac{1}{\varepsilon} \nabla \eta + \bar{u} \cdot \nabla u + u \cdot \nabla \bar{u} + \varepsilon^2 u \cdot \nabla u = 0\,, \end{aligned} \tag{1.2}$$

where $b := \omega/|\Omega|$. We shall compute the response to the wind forcing, assuming that the wind induces a **pulse at time** $t = 0$: since the wind undergoes oscillations on small spatial scales, the initial data is further assumed to depend both on x and x/ε. Typically

$$(\eta_\varepsilon, u_\varepsilon)_{|t=0} = (\eta_k(x), u_k(x)) \exp\left(i \frac{k \cdot x}{\varepsilon}\right), \tag{1.3}$$

for some $k \in \mathbb{Z}^2$. More generally we shall consider initial data which are microlocalized (in the sense of Appendix B) in some compact set of $T^*\mathbf{R}^2$.

1.4. **Local well-posedness.** The local existence of a solution to the scaled Saint-Venant Coriolis system (1.2) supplemented with initial data in the form (1.3) comes from the general theory of hyperbolic quasilinear symmetrizable systems. Defining the sound speed u_0 by

$$\eta = \left[(1 + \varepsilon^3 u_0/2)^2 - 1\right]/\varepsilon^3\,,$$

we indeed obtain that (1.2) is equivalent to

$$\varepsilon^2 \partial_t U + A(x_2, \varepsilon D) U + \varepsilon^3 Q(U) = 0\,, \qquad U = (u_0, u_1, u_2) \tag{1.4}$$

where $A(x_2, \varepsilon D_x)$ is the linear propagator

$$A(x_2, \varepsilon D) := \begin{pmatrix} \varepsilon \bar{u} \cdot \varepsilon \nabla & \varepsilon \partial_1 & \varepsilon \partial_2 \\ \varepsilon \partial_1 & \varepsilon \bar{u} \cdot \varepsilon \nabla & -b(x_2) + \varepsilon^2 \bar{u}_1' \\ \varepsilon \partial_2 & b(x_2) & \varepsilon \bar{u} \cdot \varepsilon \nabla \end{pmatrix}, \tag{1.5}$$

and $Q(U) := S_1(U)\varepsilon\partial_1 U + S_2(U)\varepsilon\partial_2 U$ with

$$S_1(U) := \begin{pmatrix} u_1 & \frac{1}{2}u_0 & 0 \\ \frac{1}{2}u_0 & u_1 & 0 \\ 0 & 0 & u_1 \end{pmatrix} \text{ and } S_2(U) := \begin{pmatrix} u_2 & 0 & \frac{1}{2}u_0 \\ 0 & u_2 & 0 \\ \frac{1}{2}u_0 & 0 & u_2 \end{pmatrix}. \tag{1.6}$$

Because of the specific form of the initial data, involving fast oscillations with respect to x, we introduce semi-classical Sobolev spaces

$$H^s_\varepsilon = \{U \in L^2 \,/\, \|U\|_{H^s_\varepsilon} < +\infty\} \text{ with } \|U\|^2_{H^s_\varepsilon} = \sum_{|k|\le s} \|(\varepsilon\nabla)^k U\|^2_{L^2}\,.$$

We shall also need in the following to define weighted semi-classical Sobolev spaces (in the spirit of [10]), adapted to the linear propagator as explained in Section 7:

$$W^s_\varepsilon := \Big\{ f \in L^2(\mathbf{R}^2) \,/\, (1-\varepsilon^2\partial_1^2)^{\frac{s}{2}}(1-\varepsilon^2\partial_2^2 + b^2(x_2))^{\frac{s}{2}} f \in L^2(\mathbf{R}^2)\Big\}. \tag{1.7}$$

A classical result based on the Sobolev embedding (see Section 7 for related results)

$$\|\varepsilon\nabla U\|_{L^\infty} \le \frac{C}{\varepsilon}\|\nabla U\|_{H^s_\varepsilon} \text{ for any } s > 1\,,$$

implies that (1.2) has a unique local solution $U_\varepsilon \in L^\infty([0,T_\varepsilon), H^{s+1}_\varepsilon)$. Note that the life span of U_ε depends a priori on ε. One of the goals of this article is to show existence on an ε-independent time interval.

2. Main results and strategy of the proofs

Most of this paper is concerned with the analysis of the solution to the linear equation

$$\varepsilon^2\partial_t V + A(x_2, \varepsilon D)V = 0\,, \qquad V = (v_0, v_1, v_2) \tag{2.1}$$

which is expected to dominate the dynamics since we consider small fluctuations. The description of the linear dynamics is provided in Theorem 1 below. The comparison between linear and nonlinear solutions is postponed to the final section of the paper (see Theorem 2).

For technical reasons we shall restrict our attention in this paper to the case of a shear flow, in the sense that $\bar{u}(x) = (\bar{u}_1(x_2), 0)$, where $\bar{u}_1$ is a smooth, compactly supported function. We shall further assume for simplicity that the zeros of $\bar{u}_1$, in the interior of its support, are **of order one**. We shall also suppose throughout the paper that b is a smooth function with a symbol-like behaviour:

$$\forall \alpha \in \mathbf{N},\ \exists C_\alpha,\ \forall y \in \mathbf{R}, \quad |b^{(\alpha)}(y)| \le C_\alpha(1 + b^2(y)), \tag{2.2}$$

and we shall further assume that $\lim_{y\to\infty} b^2(y) = \infty$, and that b has at most a finite number of critical points (that is to say points where b' vanishes). Without such an assumption one could not construct Rossby waves. We shall also suppose that the initial data is **microlocalized** (see Appendix B) in some compact set of $T^*\mathbf{R}^2$ (which we shall identify to $\mathbf{R}^4$ in the following), denoted $\mathcal{C}$ and satisfying

$$\mathcal{C} \cap \{\xi_1 = 0\} = \emptyset. \tag{2.3}$$

Thanks to this assumption, which is propagated by the linear flow, one can diagonalize the system into Rossby and Poincaré modes. Finally in order to avoid pathological trapped Rossby trajectories we shall also require that

$$\mathcal{C} \cap \Sigma = \emptyset, \tag{2.4}$$

where Σ is a codimension 1 subset of $\mathbf{R}^4$ defined in Proposition 4.6.

2.1. **Statement of the main results.** In this paragraph we shall state the two main theorems proved in this paper. The first result deals with the linear system (2.1).

Theorem 1 (The linear case)**.** *There is a submanifold Λ of $\mathbf{R}^4$, invariant under translations in the x_1-direction, such that the following properties hold.*

Let $U_{\varepsilon,0}$ be $\varepsilon-$microlocalized in a compact set $\mathcal{C}$ satisfying Assumptions (2.3)-(2.4). For any parameter $\varepsilon > 0$, denote by V_ε the associated solution to (2.1). Then for all $t \geq 0$ one can write $V_\varepsilon(t)$ as the sum of a "Rossby" vector field and a "Poincaré" vector field: $V_\varepsilon(t) = V_\varepsilon^R(t) + V_\varepsilon^P(t)$, satisfying the following properties:

(1) *There is a compact set K of $\mathbf{R}^2$ such that*
$$\forall t \geq 0, \quad \|V_\varepsilon^R(t)\|_{L^2(K)} \neq O(\varepsilon^\infty)$$
if and only if the ε-frequency set of $V_\varepsilon^R(0)$ intersects Λ.

(2) *Suppose that b^2 has only one non degenerate critical value (meaning that $(b^2)'$ only vanishes at one point, where $(b^2)''$ does not vanish). Then for any compact set Ω in $\mathbf{R}^2$, one has*
$$\forall t > 0, \quad \|V_\varepsilon^P(t)\|_{L^2(\Omega)} = O(\varepsilon^\infty).$$

In particular supposing that b^2 has only one non degenerate critical value, then there is a compact set K of $\mathbf{R}^2$ such that

$$\forall t > 0, \quad \|V_\varepsilon(t)\|_{L^2(K)} \neq O(\varepsilon^\infty)$$

if and only if the ε-frequency set of $V_\varepsilon^R(0)$ intersects Λ.

Remark 2.1. *Actually Λ corresponds to the set of initial positions and frequencies in the phase space giving rise to trapped trajectories for the Rossby hamiltonian. This will be made more precise in Section 4, where we shall prove that under some additional (non restrictive) assumptions on $\bar{u}$, Λ is of codimension one. In particular it will be shown that some of those trapped trajectories actually exhibit a singular behaviour in large times, in the sense that they converge in physical space towards a point, while the ξ_2 frequency goes to infinity. This could be interpreted like the creation of some sort of oceanic eddies.*

The result (2) is related to dispersive properties of the Poincaré hamiltonian on diffractive type times (of the type $O(1/\varepsilon^2)$), which requires some spectral analysis. Due to the assumption on b^2 one can write a rather simple proof; more general conditions could be treated, but the Bohr Sommerfeld quantization would require to decompose the phase space into various zones according to the geometry of the level sets of the Hamiltonian, which is much more technical and beyond the scope of this article. Actually in [11] *we propose a different approach, based on Mourre estimates, which allows to relax very much the assumptions on b^2 and on $\bar{u}$.*

The final section of this paper is devoted to the proof of the following theorem, which states that the very weak coupling chosen in this paper implies that nonlinear dynamics are governed by the linear equation. We also consider more generally the following weakly nonlinear system (with the notation (1.5) and (1.6)):

$$\varepsilon^2\partial_t U + A(x,\varepsilon D_x)U + \varepsilon^{3+\eta}S_1(U)\varepsilon\partial_1 U + \varepsilon^{3+\eta}S_2(U)\varepsilon\partial_2 U = 0, \quad \eta \geq 0\,. \tag{2.5}$$

The case $\eta = 0$ corresponds of course to the original system (1.4) presented in the introduction.

Theorem 2 (The nonlinear case)**.** *Let $U_{\varepsilon,0}$ be any initial data bounded in W^4_ε. Then the following results hold.*

(1) *The case $\eta = 0$:*
   (a) *There exists some $T^* > 0$ such that the initial value problem (2.5) with $\eta = 0$ has a unique solution U_ε on $[0, T^*[$ for any $\varepsilon > 0$.*
   (b) *Assume that the solution V_ε to the linear equation (2.1) satisfies*
   $$\|\varepsilon V_\varepsilon\|_{L^2([0,T^*[;L^\infty)} \to 0 \text{ as } \varepsilon \to 0.$$
   Then the solution U_ε to (2.5) with $\eta = 0$ satisfies
   $$\|U_\varepsilon - V_\varepsilon\|_{L^2} \to 0 \text{ uniformly on } [0, T^*[\text{ as } \varepsilon \to 0.$$
(2) *The case $\eta > 0$: Le $T > 0$ be fixed. Then there is $\varepsilon_0 > 0$ such that for any $\varepsilon \leq \varepsilon_0$, the equation (2.5) has a unique solution U_ε on $[0, T]$. Moreover,*
   $$\|U_\varepsilon - V_\varepsilon\|_{L^2} \to 0 \text{ uniformly on } [0, T] \text{ as } \varepsilon \to 0.$$

Remark 2.2. *Result (1b), joint with Theorem 1, implies in particular that as soon as b^2 has only one non degenerate critical value, then for positive times the energy of U_ε on any fixed compact subset is carried only by Rossby waves. The refined L^∞ estimate on the linear solution required in result (1b) should be proved by using WKB tools. For the sake of simplicity, we shall not consider such technical estimates here, all the less that we do not expect them to be enough to get an optimal result regarding the nonlinear problem (see Remark 7.1). That is the reason why we consider, in result (2), a weaker coupling still. That result implies in particular that the L^2 norm of U_ε on any fixed compact subset may remain bounded from below only if there are trapped Rossby waves, i.e. only if the ε-frequency set of the initial data does intersect Λ (with the notation of Theorem 1).*

2.2. **Some related studies.** This work follows a long tradition of mathematical studies of fast rotating fluids, following [28] and [18]; we refer for instance to [4] and [14] for a number of references. The present study concerns the case when the penalization matrix does not have constant coefficients. A first study in this type of situation may be found in [12], where a rather general penalization matrix was considered. Due to the generality of the situation, explicit computations were ruled out and no study of waves was carried out. In order to compute explicitly the modes created by the penalization matrix, various authors (see [8], [9], [10] as well as [13]) studied the betaplane approximation, in which the rotation vector depends linearly on the latitude. In that case explicit calculations may again be carried out (or some explicit commuting vector fields may be computed) and hence again one may derive envelope equations. In this paper we choose again to work with a more general rotation vector, this choice being made possible by a semi-classical setting (see the next paragraph); in particular that setting enables us diagonalize the system approximately, therefore to compute waves; note that a related study is performed by two of the authors in [5] via a purely geometric

optics approach, where no explicit diagonalization is performed (actually the initial data is strongly polarized so that only Rossby modes are present, including in the non linear setting) .

Another feature of our study is that it is a multi-scale problem, in the sense that the oscillation frequency is much bigger than the variation of the coefficients of the system. This is dealt with by using semi-classical analysis (which, compared to the previous paragraph, enables us to compute *almost commuting* vector fields although the penalization matrix no longer depends only linearly on the latitude). Such techniques are classical in geometrical optics, but in our case the additional difficulty is that the propagators are linked to different time scales: this is due to the fact that the system has eigenvalues at different scales (one is actually a subsymbol). In particular we are mostly interested in the role of the subsymbol in the dynamics, as this subsymbol is responsible for the trapping phenomenon we want to exhibit: this implies, by semi-classical analysis, the need to study the dynamical system induced by that subsymbol. On the other hand this also means that the dynamics linked to other eigenvalues must be analyzed on diffractive-type time scales, therefore much longer than that allowed by semi-classical analysis. We are able to show the dispersion of those waves by using spectral analysis and Bohr-Sommerfeld quantization.

2.3. **Organization of the paper.** The proof of Theorems 1 and 2 requires a number of steps which are described in this paragraph.

2.3.1. *Reduction to scalar propagators.* Persistent structures are related to the propagation of Rossby waves. Our first task is therefore to transform the original linear system (2.1) into three scalar equations. One is polarized on Rossby waves while the two others are polarized on Poincaré waves. This is done in Section 3 by proving some necessary conditions for the existence of those propagators. The general strategy is the following:

(1) Consider the system $A(x_2, \varepsilon D)U = i\tau U$. Take the Fourier transform in x_1, which is possible since the equation is translation invariant in x_1. Then extract from this system (by linear combinations and substitutions) a linear equation on one component u_k of U, of the type $h(x_2, \varepsilon D_2; \xi_1, \varepsilon, \tau)u_k = 0$.
(2) The symbolic equation corresponding to the PDE writes $h(x_2, \xi_2; \xi_1, \varepsilon, \tau) = 0$. It has three roots (with respect to τ) , $\tau_\pm(x_2, \xi; \varepsilon)$ (Poincaré roots) and $\tau_R(x_2, \xi; \varepsilon)$ (Rossby root). We find $\tau_\pm(x_2, \xi; \varepsilon) = \tau_\pm(x_2, \xi) + O(\varepsilon)$ and $\tau_R(x_2, \xi; \varepsilon) = \varepsilon\, \widetilde{\tau}_R(x_2, \xi) + O(\varepsilon^2)$.
(3) Those roots are not necessarily symbols. To guarantee these are indeed symbols one needs a microlocalization. Given a compact set and a truncation on that compact set χ, one can construct three operators T_j^χ (via a general theorem, stated and proved in an abstract way in Theorem 3 of Appendix A) whose principal symbols are precisely the Rossby and Poincaré symbols $\tau_\pm$ and τ_R.

2.3.2. *Trapping of Rossby waves.* Section 4 is devoted to the study of the Rossby propagator T_0, and in particular to the proof of result (1) in Theorem 1. For the time scale considered, it is easy to see that the energy propagates according to the trajectories of the semiclassical Rossby hamiltonian $\widetilde{\tau}_R$. The first step of the analysis therefore consists in studying the dynamical system giving rise to those trajectories. It turns out that the trajectories are always bounded in the x_2 direction. One is therefore reduced to studying the trajectories in the x_1

variable and in identifying the set Λ of initial data in the cotangent space giving rise to trapping in x_1. One then checks that Λ is of codimension one under some additional assumptions on $\bar{u}$, and the last step of the study consists in studying more precisely the trajectories in some specific situations, in particular in the case of the betaplane approximation.

2.3.3. *Dispersion of Poincaré waves.* The next step of our analysis of wave propagation consists in proving, in Section 5, that Poincaré waves propagate so fast that they exit from any (bounded) domain of observation on the time scale that we consider, which proves result (2) in Theorem 1. Note that, because of the very long time scaling, usual tools of semiclassical analysis cannot be applied for the Poincaré waves : we actually need deeper arguments such as the Bohr-Sommerfeld quantization to conclude.

2.3.4. *A diagonalization result.* Once the Rossby and Poincaré propagators have been well understood, we can retrace the steps followed in Section 3 to prove that the necessary conditions on the scalar propagators are sufficient. The difficulty is that the operators Π enabling one to go from the original system to the scalar equations and back (computing an approximate left inverse Q of Π at the order $O(\varepsilon^\infty)$) are only continuous on microlocalized functions; moreover the scalar propagators T_j^χ are themselves only defined on microlocalized functions. So that requires understanding the persistence of the microlocalization of the solutions to the scalar equations. That is achieved in the two previous sections, where it is proved that if the initial data is conveniently microlocalized, then for any time $t \geq 0$ one can find a compact set K and one can construct T_j^χ as in Section 3, so that the solution to the scalar equations with propagators T_j^χ is microlocalized in K (actually for Poincaré modes the microlocalization is in the variables (x_2, ξ_1, ξ_2) only, which is enough for our purpose). This enables us in Section 6 to conclude rather easily by computing explicitly the matrix principal symbols of Π and Q.

2.3.5. *The analysis of the nonlinear equation.* Section 7 is devoted to the proof that the solution of the nonlinear equation remains close to that of the linear equation. The method of proof consists first in proving the wellposedness of the nonlinear equation on a uniform time interval, by using semi-classical weighted Sobolev type spaces, whose additional feature is to be well adapted to the penalization operator $A(x_2, \varepsilon D)$: one therefore constructs a matrix-valued pseudo-differential operator which approximately commutes with $A(x_2, \varepsilon D)$. The convergence of $U_\varepsilon - V_\varepsilon$ to zero relies on a standard L^2 energy estimate on $U_\varepsilon - V_\varepsilon$, and Gronwall's lemma.

2.3.6. *Two appendixes.* In Appendix A one can find the statement and the proof of a general theorem, used in Section 3, allowing to associate to a linear evolution PDE a number of operators describing the dynamics of the equation; those operators are constructed by writing down the symbolic equation associated to the PDE and in quantizing the roots of that polynomial (in the time derivative). Appendix B finally collects a number of prerequisites on microlocal and semiclassical analysis, that are used throughout the paper.

3. Reduction to scalar propagators

3.1. **Introduction.** Let us first recall that the propagator

$$A(x,\varepsilon D)=\begin{pmatrix}\varepsilon\bar u\cdot\varepsilon\nabla & \varepsilon\partial_1 & \varepsilon\partial_2\\ \varepsilon\partial_1 & \varepsilon\bar u\cdot\varepsilon\nabla & -b+\varepsilon^2\bar u_1'\\ \varepsilon\partial_2 & b & \varepsilon\bar u\cdot\varepsilon\nabla\end{pmatrix}$$

can, in the particular case when $\bar u\equiv 0$ and $b(x_2)=\beta x_2$, be diagonalized without any error term (in particular for any finite ε), using a Fourier basis $(\exp(\frac{i}{\varepsilon}x_1\xi_1))_{\xi_1\in\mathbf{R}}$ in x_1 and a Hermite basis $(\psi_n^\varepsilon(x_2))_{n\in\mathbf{N}}$ in x_2. Precisely, the following statement is proved in [13].

Proposition 3.1 (Gallagher & Saint-Raymond,[13])**.** *For all $(\xi_1,n,j)\in\mathbf{R}\times\mathbf{N}\times\{-1,0,1\}$, denote by $\tau(\xi_1,n,j)$ the three roots (in increasing order in j) of*

$$\tau^3-(\xi_1^2+\beta\varepsilon(2n+1))\tau+\varepsilon\beta\xi_1=0. \tag{3.1}$$

Then there exists a complete family of $L^2(\mathbf{R}\times\mathbf{R},\mathbf{R}^3)$ of pseudo-eigenvectors $(\Psi^\varepsilon_{\xi_1,n,j})$ of the operator $A(x,\varepsilon D)$ (where $\bar u\equiv 0$ and $b(x_2)=\beta x_2$):

$$\forall(\xi_1,n,j)\in\mathbf{R}\times\mathbf{N}\times\{-1,0,1\},\quad A(x,\varepsilon D)\Psi^\varepsilon_{\xi_1,n,j}=i\tau(\xi_1,n,j)\Psi^\varepsilon_{\xi_1,n,j} \tag{3.2}$$

where $\Psi^\varepsilon_{\xi_1,n,j}$ can be computed in terms of the n-th Hermite function $\psi_n^\varepsilon(x_2)$ and its derivatives.

In other words, the three scalar propagators (numbered by j) can be obtained from the symbolic equation (3.1) remarking that $\beta\varepsilon(2n+1)$ is the quantization of the harmonic oscillator $-\varepsilon^2\partial_2^2+\beta^2x_2^2$. It is proved in [13] that as ξ_1 and n go to infinity

$$\tau(\xi_1,n,\pm)\sim\pm\sqrt{\xi_1^2+\beta\varepsilon(2n+1)},\quad\text{and}\quad\tau(\xi_1,n,0)\sim\frac{\varepsilon\beta\xi_1}{\xi_1^2+\beta\varepsilon(2n+1)}.$$

We are interested here in deriving a symbolic equation similar to (3.1) for a general zonal current $\bar u=(\bar u_1(x_2),0)$ and Coriolis parameter $b=b(x_2)$. The difficulty comes from the fact that the propagation of waves is governed by a matrix of differential operators with non-constant coefficients, the diagonalization of which is not a standard computation. Of course, in the semiclassical limit $\varepsilon\to 0$, we expect to get a good approximation of the propagation at leading order by considering the matrix of principal symbols

$$A_0(x,\xi):=\begin{pmatrix}0 & i\xi_1 & i\xi_2\\ i\xi_1 & 0 & -b(x_2)\\ i\xi_2 & b(x_2) & 0\end{pmatrix} \tag{3.3}$$

and by computing the scalar propagators associated to each eigenvalue

$$\tau_\pm(x_2,\xi_1,\xi_2):=\pm\sqrt{\xi_1^2+\xi_2^2+b^2(x_2)} \tag{3.4}$$

and 0. The eigenvalue 0 corresponds to the Rossby modes, whereas the two $O(1)$ eigenvalues $\pm\sqrt{\xi_1^2+\xi_2^2+b^2(x_2)}$ are the Poincaré modes. Nevertheless, this approximation is relevant only for times of order $O(\frac{1}{\varepsilon})$, and we are interested here in much longer times, of order $O(\frac{1}{\varepsilon^2})$. This means that we need to compute the next order of the expansion of the eigenvalue 0. Once that is done, we need to quantify these symbol eigenvalues to deduce scalar propagators.

3.2. **The Rossby modes.** Finding precisely the Rossby modes (up to an $O(\varepsilon^2)$ error) requires more intricate calculations than merely diagonalizing the matrix of principal symbols $A_0(x,\xi)$ given in (3.3). So let $i\tau$ be an eigenvalue of the propagator, assumed to be of the form $\tau = \varepsilon\widetilde{\tau}_R + O(\varepsilon^2)$. Explicit computations lead to the subsystem

$$\begin{pmatrix} \varepsilon\bar{u}_1\xi_1 - \tau & \xi_1 \\ \xi_1 & \varepsilon\bar{u}_1\xi_1 - \tau \end{pmatrix} \begin{pmatrix} \rho \\ u_1 \end{pmatrix} = \begin{pmatrix} i\varepsilon\partial_2 u_2 \\ -i(b - \varepsilon^2\bar{u}_1')u_2 \end{pmatrix}$$

and defining $\alpha_\tau(x_2,\xi_1) := \varepsilon\bar{u}_1\xi_1 - \tau$ and $p_R(x_2,\xi_1) = -\xi_1^2 + \alpha_\tau^2$ to the scalar equation

$$\left(\varepsilon\partial_2 p_R^{-1}\big(i\alpha_\tau\varepsilon\partial_2 + i\xi_1(b-\varepsilon^2\bar{u}_1')\big) - bp_R^{-1}(i\xi_1\varepsilon\partial_2 + i\alpha_\tau(b-\varepsilon^2\bar{u}_1')) + i\alpha_\tau\right)u_2 = 0 \tag{3.5}$$

where ξ_1 is the Fourier variable corresponding to x_1/ε. From now on we assume that ξ_1 is fixed, and is bounded away from zero (recalling Assumption (2.3)). Note that equation (3.5) makes sense because we are assuming here that $\tau = \varepsilon\widetilde{\tau}_R + O(\varepsilon^2)$, so p_R^{-1} is well defined. That would not be the case for the Poincaré modes (where $\tau = \pm\sqrt{|\xi|^2 + b^2(x_2)} + O(\varepsilon)$) so we shall use another subsystem in the next paragraph to deal with the Poincaré operators.

In order to derive a symbolic equation associated with the differential equation (3.5), we shall proceed by transforming (3.5) into a differential equation which is the left quantization (in the sense recalled in Appendix B) of a symbol, polynomial in ε and τ. This leads to a differential equation of the type

$$\mathcal{P}_2(x_2;\varepsilon,\tau)(\varepsilon\partial_2)^2u_2 + \mathcal{P}_1(x_2;\varepsilon,\tau)\varepsilon\partial_2 u_2 + \mathcal{P}_0(x_2;\varepsilon,\tau)u_2 = 0, \tag{3.6}$$

where each $\mathcal{P}_j(x_2;\varepsilon,\tau)$ is a smooth function in x_2, and has polynomial dependence in ε and in τ (precisely of degree at most 5 in ε and τ). This generalizes (3.1); one can compute in particular, using the fact that $\varepsilon\partial_2\alpha_\tau = O(\varepsilon^2)$, that

$$\mathcal{P}_2(x_2;\varepsilon,\tau) = ip_R\alpha_\tau + O(\varepsilon^2), \quad \mathcal{P}_1(x_2;\varepsilon,\tau) = O(\varepsilon^2), \quad \text{and}$$
$$\mathcal{P}_0(x_2;\varepsilon,\tau) = ip_R\left(\xi_1\varepsilon b' - (b^2+\xi_1^2)\alpha_\tau\right) + O(\varepsilon^2).$$

The differential operator appearing on the left-hand side of (3.6) is the left quantization of the following symbol:

$$h(x_2,\xi_2;\varepsilon,\tau) := -\mathcal{P}_2(x_2;\varepsilon,\tau)\xi_2^2 + i\mathcal{P}_1(x_2;\varepsilon,\tau)\xi_2 + \mathcal{P}_0(x_2;\varepsilon,\tau) \tag{3.7}$$

which belongs for each τ to $S_2(g_\tau)$ for some function g_τ of the type

$$g_\tau(x_2,\xi_2) = (1+\tau^5)\left(1 + b^2(x_2) + \xi_2^2\right)$$

recalling that $\bar{u}_1$ and $\bar{u}_1'$ are bounded from above, as well as Assumption (2.2).

Now we recall that $h(x_2,\xi_2;\varepsilon,\tau)$ is a polynomial of degree 5 in τ hence has five roots, among which 2 are actually spurious: these are of the form $\pm\xi_1 + O(\varepsilon)$, and they appear because we have multiplied the equation by a polynomial in τ which cancels at the point $\pm\xi_1$ at first order in ε. An easy computation allows to obtain the two other $O(1)$ roots, which are precisely the Poincaré roots $\pm\sqrt{|\xi|^2 + b^2(x_2)} + O(\varepsilon)$, and an asymptotic expansion allows also easily to derive the Rossby $O(\varepsilon)$ root: one finds

$$\begin{aligned} \tau_R(x_2,\xi_1,\xi_2;\varepsilon) &:= \varepsilon\widetilde{\tau}_R(x_2,\xi_1,\xi_2;\varepsilon) + O(\varepsilon^2), \quad \text{where} \\ \widetilde{\tau}_R(x_2,\xi_1,\xi_2) &:= \frac{b'(x_2)\xi_1}{\xi_1^2+\xi_2^2+b^2(x_2)} + \xi_1\bar{u}_1(x_2). \end{aligned} \tag{3.8}$$

Now that the root τ_R has been computed, our next task is to prove the existence of the Rossby propagator $T_R = \varepsilon\widetilde{T}_R$ whose principal symbol is precisely $\varepsilon\widetilde{\tau}_R$. Actually this result is a direct

consequence of Theorem 3 stated and proved in Appendix A: with the notation of Theorem 3, one has $\nu = 1$ and $\partial_\tau h_0(x,\xi,0) = -i\xi_1^2(\xi_1^2+\xi_2^2+b^2(x_2))$. In the following statement, the time variable s is defined as $s = t/\varepsilon^2$.

Proposition 3.2 (The Rossby propagator)**.** *Let $\widetilde{\tau}_R$ be the symbol defined in (3.8). Then for any compact set $\mathcal{K}$ satisfying Assumptions (2.3,2.4) there exists a formally self-adjoint pseudo-differential operator $\widetilde{T}_R$ of principal symbol $\widetilde{\tau}_R$ such that if φ_R is microlocalized in $\mathcal{K}$ and solves*

$$\partial_s \varphi_R = i\varepsilon \widetilde{T}_R \varphi_R, \tag{3.9}$$

then

$$\mathcal{P}_2(x_2;\varepsilon,\partial_s)(\varepsilon\partial_2)^2\varphi_R + \mathcal{P}_1(x_2;\varepsilon,\partial_s)(\varepsilon\partial_2)\varphi_R + \mathcal{P}_0(x_2;\varepsilon,\partial_s)\varphi_R = O(\varepsilon^\infty).$$

Definition 3.3 (The Rossby operator)**.** *We shall call Π_R the Rossby operator defined by*

$$\Pi_R := \begin{pmatrix} P_R^{-1}\big(i(-i\varepsilon^2\bar{u}_1\partial_1 - \varepsilon\widetilde{T}_R)\varepsilon\partial_2 + \varepsilon\partial_1(b-\varepsilon^2\bar{u}_1')\big) \\ -P_R^{-1}\big(-\varepsilon^2\partial_{12} + i(-i\varepsilon^2\bar{u}_1\partial_1 - \varepsilon\widetilde{T}_R)(b-\varepsilon^2\bar{u}_1')\big) \\ \mathrm{Id} \end{pmatrix},$$

where $P_R := \varepsilon^2\partial_1^2 - (i\varepsilon^2\bar{u}_1\partial_1 + \varepsilon\widetilde{T}_R)^2$.

Remark 3.4. (1) *Notice that Π_R is well defined since the principal symbol of P_R is bounded from below (see Appendix B).*

(2) *Proposition 4.6 shows that if $\varphi_{R|t=0}$ is microlocalized in a compact set $\mathcal{K}_0$ satisfying (2.4), then the solution to (3.9) is microlocalized for all $t \geq 0$ in a compact set $\mathcal{K}_t$.*

(3) *The above computations allow to formally recover the original shallow-water equation, up to $O(\varepsilon^\infty)$. Indeed retracing the steps which enabled us above to derive equation (3.6) shows that if φ^0 is a smooth function conveniently microlocalized, and if φ solves*

$$\varepsilon\partial_t\varphi = i\widetilde{T}_R\varphi, \quad \varphi_{|t=0} = \varphi^0,$$

then the vector field $U := \Pi_R\varphi$ satisfies (2.1) up to $O(\varepsilon^\infty)$. This property will be made rigorous in Section 6.

3.3. **The Poincaré modes.** In this paragraph we shall follow the method used above in the case of Rossby modes to infer Poincaré propagators $T_\pm$ and operators $\Pi_\pm$. Actually one cannot use precisely the same method since the symbol $(\varepsilon\bar{u}_1\xi_1 - \tau)^2 - \xi_1^2$ may vanish when $\tau = \tau_\pm + O(\varepsilon)$. So we shall instead consider the subsystem

$$\begin{pmatrix} \varepsilon\bar{u}_1\xi_1 - \tau & -i\varepsilon\partial_2 \\ -i\varepsilon\partial_2 & \varepsilon\bar{u}_1\xi_1 - \tau \end{pmatrix}\begin{pmatrix} \rho \\ u_2 \end{pmatrix} = \begin{pmatrix} -\xi_1 u_1 \\ ibu_1 \end{pmatrix}$$

and the scalar equation

$$\Big(\xi_1 p_P^{-1}(-\alpha_\tau\xi_1 - \varepsilon\partial_2 b) + \alpha_\tau + (b-\varepsilon^2\bar{u}_1')p_P^{-1}(\varepsilon\partial_2 - \alpha_\tau b)\Big)u_1 = 0, \tag{3.10}$$

where as before $\alpha_\tau(x_2,\xi_1) := \varepsilon\bar{u}_1\xi_1 - \tau$, and $p_P(x_2,\xi_1) := \varepsilon^2\partial_2^2 + \alpha_\tau^2$. We notice here that α_τ^{-1} is well defined when $\tau = \tau_\pm$ since $\tau_\pm$ is bounded away from zero by the assumption on ξ_1, and the same goes for $\tau_\pm^2 - \xi_2^2$ so p_P^{-1} is also well defined. Then it remains to follow the steps of Paragraph 3.2 to obtain a new scalar PDE of the same type as (3.6), as well as a symbol equation of the type (3.7):

$$\tilde{h}(x_2,\xi_2;\varepsilon,\tau) := -\tilde{\mathcal{P}}_2(x_2;\varepsilon,\tau)\xi_2^2 + i\tilde{\mathcal{P}}_1(x_2;\varepsilon,\tau)\xi_2 + \tilde{\mathcal{P}}_0(x_2;\varepsilon,\tau). \tag{3.11}$$

Of course $\tau_\pm$ are roots of that equation up to $O(\varepsilon)$, and $\varepsilon\tau_R$ is a root up to $O(\varepsilon^2)$. Then the application of Theorem 3 implies a similar result to Proposition 3.2, noticing that with the notation of Theorem 3, $\nu = 0$ and $\partial_\tau h_0(x,\xi,\tau_\pm) = -2(\xi_1^2+\xi_2^2+b^2(x_2))$.

Proposition 3.5 (The Poincaré propagator)**.** *Let $\tau_\pm$ be the symbol defined in (3.4). Consider a compact set $\mathcal{K}_P \subset \mathbf{R}^* \times T^*\mathbf{R}$. Then there exists formally self-adjoint pseudo-differential operators $T_\pm$ of principal symbols $\tau_\pm$ such that if $\varphi_\pm$ solves $\partial_s\varphi_\pm = iT_\pm\varphi_\pm$, and $\varphi_\pm$ is microlocalized in $\mathbf{R}\times\mathcal{K}_P$, then*

$$\tilde{\mathcal{P}}_2(x_2;\varepsilon,\partial_s)(\varepsilon\partial_2)^2\varphi_\pm + \tilde{\mathcal{P}}_1(x_2;\varepsilon,\partial_s)(\varepsilon\partial_2)\varphi_\pm + \tilde{\mathcal{P}}_0(x_2;\varepsilon,\partial_s)\varphi_\pm = O(\varepsilon^\infty).$$

Definition 3.6 (The Poincaré operator)**.** *We shall call $\Pi_\pm$ the Poincaré operator defined by*

$$\Pi_\pm := \begin{pmatrix} P_\pm^{-1}(i(-i\varepsilon^2\bar{u}_1\partial_1 - T_\pm)\varepsilon\partial_1 + \varepsilon\partial_2 b) \\ \mathrm{Id} \\ P_\pm^{-1}(i(-i\varepsilon^2\bar{u}_1\partial_1 - T_\pm)b - \varepsilon\partial_1\varepsilon\partial_2) \end{pmatrix}, \quad \text{where} \quad P_\pm := (i\varepsilon^2\bar{u}_1\partial_1 + T_\pm)^2 + \varepsilon^2\partial_2^2.$$

Remark 3.7. *As in Remark 3.4, one sees formally that if φ^0 is a smooth function conveniently microlocalized, and if φ solves*

$$\varepsilon^2\partial_t\varphi = iT_\pm\varphi, \quad \varphi_{|t=0} = \varphi^0,$$

then the vector field $U := \Pi_\pm\varphi$ satisfies (2.1) up to $O(\varepsilon^\infty)$. This property will be made rigorous in Section 6.

4. Study of the Rossby waves

4.1. **The dynamical system.** For the time scale considered here, the propagation of energy by Rossby waves is given by the transport equation (see Appendix B)

$$\partial_t f + \{\widetilde{\tau}_R, f\} = 0$$

where $\widetilde{\tau}_R$ is the principal symbol of the Rossby mode computed in (3.8):

$$\widetilde{\tau}_R(\xi_1,x_2,\xi_2) = \frac{b'(x_2)\xi_1}{\xi_1^2+\xi_2^2+b^2(x_2)} + \bar{u}_1(x_2)\,\xi_1\,. \tag{4.1}$$

As $\widetilde{\tau}_R$ is a smooth function of (x_2,ξ_1,ξ_2), the energy is propagated along the bicharacteristics, i.e. along the integral curves of the following system of ODEs:

$$\begin{cases} \dot{x}^t = \nabla_\xi\widetilde{\tau}_R(\xi_1^t,x_2^t,\xi_2^t), & x^0 = (x_1^0,x_2^0) \\ \dot{\xi}^t = -\nabla_x\widetilde{\tau}_R(\xi_1^t,x_2^t,\xi_2^t), & \xi^0 = (\xi_1^0,\xi_2^0). \end{cases}$$

Since the condition (2.3) avoids the set $\{\xi_1 = 0\}$, we can suppose that $\xi_1^0 \neq 0$. Moreover since $\widetilde{\tau}_R$ does not depend on x_1, we find $\xi_1^t \equiv \xi_1^0$. The ODE to be studied is therefore

$$\begin{cases} \dot{x}_1^t &= \bar{u}_1(x_2^t) + \dfrac{b'(x_2^t)\,(-\xi_1^2+{\xi_2^t}^2+b^2(x_2^t))}{(\xi_1^2+{\xi_2^t}^2+b^2(x_2^t))^2} \\ \dot{x}_2^t &= \dfrac{-2b'(x_2^t)\,\xi_1\xi_2^t}{(\xi_1^2+{\xi_2^t}^2+b^2(x_2^t))^2} \\ \dot{\xi}_2^t &= -\bar{u}_1'(x_2^t)\xi_1 + \dfrac{2b(x_2^t)b'(x_2^t)^2\,\xi_1}{(\xi_1^2+{\xi_2^t}^2+b^2(x_2^t))^2} - \dfrac{b''(x_2^t)\,\xi_1}{\xi_1^2+{\xi_2^t}^2+b^2(x_2^t)}\,. \end{cases} \tag{4.2}$$

Due to the assumptions on $\bar{u}_1$ and on b, the map $(x,\xi)\mapsto(\nabla_\xi\widetilde{\tau}_R, -\nabla_x\widetilde{\tau}_R)(x,\xi)$ is bounded, so the integral curves are globally defined in time. The strategy to study their qualitative

behaviours is to first (in Section 4.2) consider the motion in the reduced phase space $(x_2, \xi_2) \in \mathbf{R}^2$ and then (in Section 4.3) to study the motion in the x_1 direction.

4.2. **Trajectories in the reduced (x_2, ξ_2) phase space.** In this section we study the trajectories in the reduced (x_2, ξ_2) phase space. We shall denote $\xi_1 := \xi_1^0$.

4.2.1. *Energy surfaces.* Since the Hamiltonian $\widetilde{\tau}_R$ and ξ_1 are conserved along any trajectory, trajectories are submanifolds of

$$\mathcal{E}_{\tau,\xi_1} := \big\{(x_2, \xi_2) \in \mathbf{R}^2\,; \widetilde{\tau}_R(\xi_1, x_2, \xi_2) = \tau\big\}.$$

In the following we shall note for any energy τ and any $\xi_1 \in \mathbf{R}^*$

$$V_{\tau,\xi_1}(x_2) := \frac{b'(x_2)\xi_1}{\tau - \bar{u}_1(x_2)\xi_1} - \xi_1^2 - b^2(x_2),$$

so that if $\mathcal{D} := \big\{x_2 / V_{\tau,\xi_1}(x_2) \geq 0\big\}$, then $\mathcal{E}_{\tau,\xi_1} = \Big\{\big(x_2, \pm\sqrt{V_{\tau,\xi_1}(x_2)}\,\big),\ x_2 \in \mathcal{D}\Big\}$. Note that $V_{\tau,\xi_1}(x_2^t)$ becomes singular if x_2^t reaches a point x_2 such that $\tau = \bar{u}_1(x_2)\xi_1$.

Proposition 4.1. *The projection of $\mathcal{E}_{\tau,\xi_1}$ on the x_2-axis is bounded.*

Proof. We recall that on $\mathcal{E}_{\tau,\xi_1}$ we have

$$\frac{b'(x_2)\xi_1}{\xi_1^2 + \xi_2^2 + b^2(x_2)} + \bar{u}_1(x_2)\xi_1 = \tau.$$

Suppose the trajectory in x_2 is not bounded, then in particular it escapes the support of $\bar{u}_1$. In the case when $\tau \neq 0$, letting $|x_2|$ go to infinity yields a contradiction due to the assumptions on b. In the case $\tau = 0$, for x_2 out of the support of $\bar{u}_1$ we have

$$\frac{b'(x_2)\xi_1}{\xi_1^2 + \xi_2^2 + b^2(x_2)} = 0$$

and the only possibility is for x_2 to be fixed on a zero point of b' (hence in particular does not go to infinity). □

Proposition 4.1 shows that to prove that some trajectories are trapped in physical space, it suffices to study their behaviour in the x_1 direction. However before doing so, let us prepare that study by classifying the trajectories in the reduced phase space. Up to a change of parameter, namely expressing time t as a function of x_2

$$dt = \pm\frac{b'(x_2)\xi_1}{(\tau - \xi_1\bar{u}_1(x_2))^2\sqrt{V_{\tau,\xi_1}(x_2)}}dx_2\,,$$

which is justified locally, and will give the convenient global behaviour by suitable gluing, we are brought back to the study of the hamiltonian system $\xi_2^2 - V_{\tau,\xi_1}(x_2)$ describing the motion of a particle in the potential $-V_{\tau,\xi_1}$.

For smooth potentials V such that $V(x) \to -\infty$ as $|x| \to \infty$, the possible behaviours of such a system are well-known and the trajectories are usually classified as follows (see for instance [1, 2, 3, 20, 22]): periodic orbits, fixed points, homoclinic and heteroclinic orbits connecting unstable fixed points. Here the situation is more complex insofar as V_{τ,ξ_1} admits singularities. We shall classify the trajectories according to their motion in the x_2 variable.

4.2.2. *Periodic trajectories.* These correspond to the case when there exists $[x_{min}, x_{max}]$ in $\mathbf{R}$ with $x_{min} \neq x_{max}$, containing x_2^0 such that

- V_{τ,ξ_1} has no singularity and does not vanish on $]x_{min}, x_{max}[$;
- $V_{\tau,\xi_1}(x_{min}) = V_{\tau,\xi_1}(x_{max}) = 0$;
- the points x_{min} and x_{max} are reached in finite time.

The extremal points x_{min} and x_{max} are then turning points, meaning that the motion is periodic. The fact that x_{min} and x_{max} are reached in finite time is equivalent to $V'_{\tau,\xi_1}(x_{min}) \neq 0$ and $V'_{\tau,\xi_1}(x_{max}) \neq 0$. Indeed if $V'_{\tau,\xi_1}(x_{max}) = 0$ (resp. $V'_{\tau,\xi_1}(x_{min}) = 0$), then $(x_{max}, 0)$ (resp $(x_{min}, 0)$) is a fixed point, which contradicts the uniqueness given by the Cauchy-Lipschitz theorem. And conversely, if $V'_{\tau,\xi_1}(x_{max}) \neq 0$ (resp. $V'_{\tau,\xi_1}(x_{min}) \neq 0$), an asymptotic expansion in the vicinity of x_{max} (resp. x_{min}) shows that the extremal point is reached in finite time.

Definition 4.2. *We will denote by $\mathcal{P}$ the subset of the phase space $T^*\mathbf{R}^2$ consisting of initial data corresponding to periodic motions along x_2.*

A rather simple continuity argument allows to prove that $\mathcal{P}$ is an open subset of $\mathbf{R} \times \mathbf{R}^* \times \mathbf{R}^2$. Denote indeed by $(\tilde{x}_1^0, \tilde{\xi}_1, \tilde{x}_2^0, \tilde{\xi}_2^0)$ any point of $\mathcal{P}$, and by $\tilde{x}_{min}$ and $\tilde{x}_{max}$ the extremal points of the corresponding (periodic) trajectory along x_2. As $V'_{\tilde{\tau},\tilde{\xi}_1}(\tilde{x}_{max}) \neq 0$ and V_{τ,ξ_1} is a smooth function of τ and ξ_1 outside from the closed subset of singularity points, the implicit function theorem gives the existence of a neighborhood of $(\tilde{\xi}_1, \tilde{\tau})$ such that there exists a unique x_{max} which satisfies $V_{\tau,\xi_1}(x_{max}) = 0$. Furthermore x_{max} depends continuously on τ and ξ_1, in particular $V'_{\tau,\xi_1}(x_{max}) \neq 0$. Using the same arguments to build a suitable x_{min}, we finally obtain that there exists a neighborhood of $(\tilde{x}_1^0, \tilde{\xi}_1, \tilde{x}_2^0, \tilde{\xi}_2^0)$ for which the motion along x_2 is a non degenerate periodic motion. Moreover, x_{min}, x_{max} and also ξ_{min}, ξ_{max} and the period T depend continuously on the initial data.

Fixed points correspond to the degenerate case when $x_{min} = x_{max} = x_2^0$, which implies that either $\xi_2^0 = 0$ or $b'(x_2^0) = 0$. The latter case is completely characterized by the condition $b'(x_2^0) = 0$, so let us focus on the case when $b'(x_2^0) \neq 0$. Fixed points correspond then to local extrema of V_{τ,ξ_1}. They can be either stable or unstable depending on the sign of V''_{τ,ξ_1}.

Stable fixed points are obtained as a limit of periodic orbits when the period $T \to 0$, whereas unstable fixed points are obtained in the limit $T \to \infty$ as explained below.

Stopping trajectories belong to the same energy surfaces as unstable fixed points and reach some unstable fixed point in infinite time : they correspond to the case when there exists an interval $[x_{min}, x_{max}]$ of $\mathbf{R}$ containing x_2^0 such that

- V_{τ,ξ_1} has no singularity and does not vanish on $]x_{min}, x_{max}[$;
- x_{min} and x_{max} are either zeros or singularities of V_{τ,ξ_1};
- as $t \to \infty$, $x_2^t \to x_2^\infty$ such that $V'_{\tau,\xi_1}(x_2^\infty) = 0$ or $b'(x_2^\infty) = 0$.

They are in some sense also a degenerate version of periodic trajectories since for arbitrarily close initial data, one can obtain periodic orbits.

Definition 4.3. *We will denote by $\delta\mathcal{P}$ the subset of the phase space $T^*\mathbf{R}^2$ consisting in initial data corresponding to fixed points and stopping motions along x_2.*

Using the characterization of the energy surfaces which carry such pathological motions, we can prove that $\delta\mathcal{P}$ is a codimension 1 subset of the phase space. We first consider the energy surfaces containing a fixed point $\tilde{x}_2$ such that $b'(\tilde{x}_2) = 0$. We then have $\tau = \xi_1 \bar{u}_1(\tilde{x}_2)$. The corresponding set of initial data

$$\{(x_1^0, \xi_1, x_2^0, \xi_2^0) \,/\, \widetilde{\tau}_R(\xi_1, x_2^0, \xi_2^0) = \xi_1 \bar{u}_1(\tilde{x}_2)\} \text{ is of codimension 1.}$$

As we have assumed that b has only a finite number of critical points, the union of these sets is still of codimension 1. We then consider the energy surfaces containing a fixed point $(\tilde{x}_1, \tilde{\xi}_1, \tilde{x}_2, 0)$ such that $V_{\tilde{\tau},\tilde{\xi}_1}(\tilde{x}_2) = V'_{\tilde{\tau},\tilde{\xi}_1}(\tilde{x}_2) = 0$. We then have

$$-\bar{u}_1'(\tilde{x}_2) + \frac{2b(\tilde{x}_2)b'(\tilde{x}_2)}{(\tilde{\xi}_1^2 + b^2(\tilde{x}_2))^2} - \frac{b''(\tilde{x}_2)}{\tilde{\xi}_1^2 + b^2(\tilde{x}_2)} = 0.$$

For each $\tilde{x}_2$, there are at most two values of $\tilde{\xi}_1^2$ such that the previous quantity vanishes. We therefore deduce that

$$\Big\{(x_1^0, \xi_1, x_2^0, \xi_2^0) \,/\, \widetilde{\tau}_R(\xi_1, x_2^0, \xi_2^0) = \frac{b'(\tilde{x}_2)\xi_1}{\xi_1^2 + b^2(\tilde{x}_2)} + \bar{u}_1(\tilde{x}_2)\,\xi_1 \quad \text{and}$$
$$-\bar{u}_1'(\tilde{x}_2) + \frac{2b(\tilde{x}_2)b'(\tilde{x}_2)}{(\tilde{\xi}_1^2 + b^2(\tilde{x}_2))^2} - \frac{b''(\tilde{x}_2)}{\tilde{\xi}_1^2 + b^2(\tilde{x}_2)} = 0\Big\} \text{ is of codimension 1.}$$

It consists indeed of at most eight manifolds, each one of them parametrized by the real parameter $\tilde{x}_2$. In the sequel, we shall avoid these pathological motions assuming that the initial data is microlocalized outside $\delta\mathcal{P}$.

4.2.3. *Asymptotic trajectories.* These correspond to the case when there exists an interval $[x_{min}, x_{max}]$ of $\mathbf{R}$ containing x_2^0 such that

- V_{τ,ξ_1} has no singularity and does not vanish on $]x_{min}, x_{max}[$;
- x_{min} and x_{max} are either zeros or singularities of V_{τ,ξ_1};
- as $t \to \infty$, $x_2^t \to x_2^\infty$ where $x_2^\infty \in \{x_{min}, x_{max}\}$ is a pole of multiplicity 1 of V_{τ,ξ_1}. For the sake of simplicity, we further impose that $b'(x_2^\infty) \neq 0$.

This situation therefore corresponds to a motion which is not periodic.

Depending on the sign of ξ_1, an asymptotic trajectory will either encounter a turning point and then converge asymptotically to the singular point, or converge monotonically to the limiting point. As x_2^∞ is such that $\bar{u}_1(x_2^\infty) = \tau/\xi_1$, one has

- either $\tau \neq 0$ and x_2^∞ belongs to the support of $\bar{u}_1$,
- or $\tau = 0$ and $\bar{u}_1(x_2^0) = -b'(x_2^0)/(\xi_1^2 + \xi_2^2 + b^2(x_2^0)) \neq 0$, meaning that x_2^0 belongs to the support of $\bar{u}_1$. Therefore, x_2^∞ is either $\min\{y > x_2^0 \,/\, \bar{u}_1(y) = 0\}$ or $\max\{y < x_2^0 \,/\, \bar{u}_1(y) = 0\}$, in particular x_2^∞ belongs to the support of $\bar{u}_1$.

Definition 4.4. *We will denote by $\mathcal{A}$ the subset of the phase space consisting of initial data corresponding to asymptotic motions along x_2.*

The same kind of arguments as in the previous paragraph allow to prove that $\mathcal{A}$ is an open subset of the phase space $\mathbf{R} \times \mathbf{R}^* \times \mathbf{R}^2$. Consider indeed some $(\tilde{x}_1^0, \tilde{\xi}_1, \tilde{x}_2^0, \tilde{\xi}_2^0) \in \mathcal{A}$, and the corresponding asymptotic point $\tilde{x}_2^\infty$. As x_2^∞ is a pole of multiplicity 1 of $V_{\tilde{\tau},\tilde{\xi}_1}$, the implicit

function theorem shows that V_{τ,ξ_1}^{-1} admits locally a unique zero, which depends continuously on τ and ξ_1. Using further the continuity of the possible turning point, we get that $\mathcal{A}$ contains a neighborhood of $(\tilde{x}_1^0, \tilde{\xi}_1, \tilde{x}_2^0, \tilde{\xi}_2^0)$. Moreover, we can obtain bounds on the expansion (with respect to time) of any compact subset of $\mathcal{A}$. Here we will focus on the growth of ξ_2^t, and proves that it depends continuously on the initial data in $\mathcal{A}$. Without loss of generality, we can consider the case when the asymptotic point is x_{max}. Then we recall that

$$\lim_{t\to+\infty} \xi_2^t = \infty \quad \text{and} \quad \lim_{t\to+\infty} x_2^t = x_2^\infty, \quad \text{with} \quad \xi_1 \bar{u}_1(x_2^\infty) = \tau.$$

As x tends to x_2^∞, we have (recalling that $b'(x_2^\infty) \neq 0$)

$$V_\tau(x) \sim -\frac{b'(x_2^\infty)}{\bar{u}_1'(x_2^\infty)}(x - x_2^\infty)^{-1}\,.$$

This implies that

$$|\xi_2^t|^2 \sim -\frac{b'(x_2^\infty)}{\bar{u}_1'(x_2^\infty)}(x - x_2^\infty)^{-1} \quad \text{and} \quad \dot{x}_2^t \sim -2b'(x_2^\infty)\xi_1 \frac{\xi_2^t}{|\xi_2^t|^4} \sim 2\frac{|\bar{u}_1'(x_2^\infty)|^{3/2}|\xi_1|}{|b'(x_2^\infty|^{1/2}}(x_2^\infty - x)^{3/2}\,.$$

By integration, we get

$$x_2^t \sim x_2^\infty + C_1\, t^{-2}, \qquad \xi_2^t \sim C_2 t$$

where C_1 and C_2 depend continuously on x_2^∞, and consequently on the initial data.

Singular trajectories are a degenerate version of the asymptotic trajectories above: they correspond to the case when there exists an interval $[x_{min}, x_{max}]$ of $\mathbf{R}$ containing x_2^0 such that

- V_{τ,ξ_1} has no singularity and does not vanish on $]x_{min}, x_{max}[$;
- x_{min} and x_{max} are either zeros or singularities of V_τ;
- as $t \to \infty$, $x_2^t \to x_2^\infty$ where $x_2^\infty \in \{x_{min}, x_{max}\}$ is either a singularity of order greater than 1 of V_{τ,ξ_1}, or a singularity which is also a zero of b'.

As previously, it is easy to check that x_2^∞ is necessarily in the support of $\bar{u}_1$. Furthermore, we have $\bar{u}_1'(x_2^\infty) = 0$. Indeed, in the vicinity of x_2^∞, we have

$$V_{\tau,\xi_1}(y) \sim \frac{b'(x_2^\infty)\xi_1}{\bar{u}(x_2^\infty) - \bar{u}(y)}$$

which has a singularity of order greater than 1, or a singularity which is also a zero of b' if and only if $\bar{u}'(x_2^\infty) = 0$.

Definition 4.5. *We will denote by $\delta\mathcal{A}$ the subset of the phase space consisting of initial data corresponding to singular motions along x_2.*

The previous condition $\bar{u}_1'(x_2^\infty) = 0$ shows that $\delta\mathcal{A}$ is included in the union of energy surfaces

$$\mathcal{E}_{\tau,\xi_1} \text{ with } \tau = \bar{u}_1(y) \text{ for some } y \in supp(\bar{u}_1) \text{ such that } \bar{u}_1'(y) = 0\,.$$

As $\bar{u}_1'$ has only a finite number of zeros in the support of $\bar{u}_1$, this implies that $\delta\mathcal{A}$ is a codimension 1 subset of the phase space.

Gathering all the previous results together, we obtain the following

Proposition 4.6. *The phase space* $\mathbf{R} \times \mathbf{R}^* \times \mathbf{R}^2$ *admits the following partition*

$$\mathbf{R} \times \mathbf{R}^* \times \mathbf{R}^2 = \mathcal{P} \cup \mathcal{A} \cup \Sigma$$

where $\mathcal{P}$ *and* $\mathcal{A}$ *are the open sets of initial data giving rise respectively to periodic motions and asymptotic motions along* x_2*, and* $\Sigma := \delta\mathcal{P} \cup \delta\mathcal{A}$ *is the codimension 1 set of initial data giving rise to pathological motions along* x_2*.*

For any compact set $K \subset \mathcal{A} \cup \mathcal{P}$*, and for any time* $T > 0$*, we further have a uniform bound on the image of* K *by the flow up to time* T*.*

Proof. The first statement just tells us that all trajectories belong to one of the four categories described above. Indeed, for any initial data $(x_1^0, \xi_1, x_2^0, \xi_2^0)$, one has

$$V_{\tau,\xi_1}(x_2^0) = (\xi_2^0)^2 \geq 0, \text{ and } V_{\tau,\xi_1}(x_2) \to -\infty \text{ as } |x_2| \to \infty,$$

so that there exists an interval $[x_{min}, x_{max}]$ of $\mathbf{R}$ containing x_2^0 such that V_{τ,ξ_1} has no singularity and does not vanish on $]x_{min}, x_{max}[$, and x_{min} and x_{max} are either zeros or singularities of V_{τ,ξ_1}. The second statement is then a simple corollary of the continuity results established on $\mathcal{P}$ and $\mathcal{A}$. □

4.3. **Analysis of the trajectories in the x_1 direction: trapping phenomenon.** Proposition 4.1 states that the trajectories are always bounded in the x_2 variable, so it remains to study the x_1 variable. Our aim is to find a set $\Lambda \subset \mathbf{R} \times \mathbf{R}^* \times \mathbf{R}^2$ such that any initial data $(x_1^0, \xi_1, x_2^0, \xi_2^0)$ in Λ gives rise to a trapped trajectory, meaning that

$$\int_0^t \dot{x}_1^s ds \text{ is uniformly bounded for } t \in \mathbf{R}^+ .$$

4.3.1. *The criterion of capture.* Let us prove the following result.

Proposition 4.7. *A necessary and sufficient condition for a trajectory with initial data in* $\mathcal{A} \cup \mathcal{P}$ *to be trapped is*

$$\lim_{t \to T} \frac{1}{t} \int_0^t \dot{x}_1^s ds = 0,$$

where T *denotes the (finite) period of the motion along* x_2 *in the periodic case, and* $T = +\infty$ *in the asymptotic case.*

Proof. We will study separately the different situations described in the previous section, namely the case of periodic and asymptotic trajectories in (x_2, ξ_2).

• *In the case of a periodic motion* in (x_2, ξ_2) of period $T > 0$, the function $\dot{x}_1^t$ is also periodic, with the same period. Writing

$$x_1^t = x_1^0 + \int_0^t \left(\dot{x}_1^s - \frac{1}{T} \int_0^T \dot{x}_1^{s'} \, ds' \right) ds + \frac{t}{T} \int_0^T \dot{x}_1^s \, ds$$

we see that depending on the average of $\dot{x}_1^t$ over $[0, T]$, x_1^t is either a periodic function, or the sum of a periodic function and a linear function. It follows that trapped trajectories are characterized by the criterion $\int_0^T \dot{x}_1^t dt = 0$. Note that, depending on the period T, the trajectory can explore a domain in x_1 the size of which may be very large. Nevertheless the

continuity statement in Proposition 4.6 shows that we have a uniform bound on this size on any compact subset of $\mathcal{P}$.

• For asymptotic motions, we need to check that

$$\dot{x}_1^t - \bar{u}_1(x_2^\infty)t \text{ is integrable at infinity.}$$

We have indeed

$$\dot{x}_1^t = \bar{u}_1(x_2^t) + \frac{b'(x_2^t)(-\xi_1^2 + {\xi_2^t}^2 + b^2(x_2^t))}{(\xi_1^2 + {\xi_2^t}^2 + b^2(x_2^t))^2},$$

which, together with the asymptotic expansions of x_2^t and ξ_2^t obtained in the previous section, implies that

$$\dot{x}_1^t = \bar{u}_1(x_2^\infty) + O(t^{-2}).$$

It is then clear that the trajectory is trapped if and only if

$$\bar{u}_1(x_2^\infty) = \lim_{t\to\infty} \frac{1}{t}\int_0^t \dot{x}_1^s ds = 0\,.$$

Using again the continuity statement in Proposition 4.6, we also get a uniform bound on the size of the time evolution of any compact subset of $\mathcal{A}$. □

Remark 4.8. *Note that, in the case of a singular asymptotic motion along x_2, the criterion of capture is equivalent to $\bar{u}(x_2^\infty) = \tau = 0$. In the case of periodic motions along x_2, this criterion - even more complicated - can also be expressed in terms of the initial data. The map x_2^t is indeed a smooth bijection from a time interval $]t_0, t_0 + \frac{T}{2}[$ to $]x_{min}, x_{max}[$, so*

$$\int_0^T \dot{x}_1^t\, dt \;=\; 2\int_{t_0}^{t_0+T/2} \dot{x}_1^t\, dt \;=\; 2\int_{x_{min}}^{x_{max}} \dot{x_1}^t_{|\,x_2^t=x_2}\; (\dot{x}_2^t)^{-1}_{|\,x_2^t=x_2}\; dx_2\,.$$

The trajectory is therefore trapped if and only if

$$\int_{x_{min}}^{x_{max}} \Big[\xi_1^2 - \frac{(\tau/\xi_1)\, b'(y)}{2\big((\tau/\xi_1) - \bar{u}_1(y)\big)^2}\Big]\, V_{\tau,\xi_1}(y)^{-1/2}\, dy = 0 \tag{4.3}$$

(notice that since we are only looking for a criterion for the function to vanish, we can replace ξ_2 by $\sqrt{V_{\tau,\xi_1}}$ without discussing the sign). Such a formula would be useful to investigate numerically the initial data giving rise to trapped trajectories. We shall also use it when investigating in more detail the case of the betaplane approximation, in Paragraph 4.4.

4.3.2. *Exhibiting a subset of Λ of codimension 1.* Let us prove the following proposition.

Proposition 4.9. *Let $\bar{u}_1 \in C_c^\infty(\mathbf{R})$ be a function which is not identically positive, with some zero of finite multiplicity. Then the set Λ_{sing} consisting of data in Λ giving rise to singular and trapped trajectories, is nonempty. It contains a submanifold of $\mathbf{R}\times\mathbf{R}^*\times\mathbf{R}^2$ which is of codimension 1.*

Proof. Without loss of generality, we can assume that there exist $y_1 < y_2$ where $\bar{u}_1$ vanishes, with $\bar{u}_1'(y_2) > 0$ and $\bar{u}_1(y) < 0$ on $]y_1, y_2[$. As we want to study trapped asymptotic trajectories, we will restrict our attention to the case when $\tau = 0$, which is a necessary condition for asymptotic trajectories to be trapped. Extremal points of the trajectories are then defined in terms of the function

$$V_{0,\xi_1}(y) = -\frac{b'(y)}{\bar{u}_1(y)} - b(y)^2 - \xi_1^2\,.$$

More precisely, if we introduce the auxiliary function

$$\varrho(y) := -\frac{b'(y)}{\bar{u}_1(y)} - b(y)^2\,, \qquad y \in]y_1, y_2[\,,$$

we obtain turning (or stopping) points y_s if $\rho(y_s) = \xi_1^2$, and singular points y_s if $\lim_{y\to y_s} \rho(y) = +\infty$ or equivalently $\bar{u}_1(y_s) = 0$. By definition of y_1 and y_2, one has

$$\lim_{y\to y_1+} \varrho(y) = +\infty \quad \text{and} \quad \lim_{y\to y_2-} \varrho(y) = +\infty.$$

Let us define $N := \max\big(0\,; \inf_{y\in]y_1,y_2[} \varrho(y)\big) \in \mathbf{R}_+$. For ξ_1 such that $\xi_1^2 \geq N$, we then define

$$h(\xi_1) := \sup\big\{\, y \in]-\infty, y_2[\,;\ \varrho(y) \leq \xi_1^2 \,\big\} \in]y_1, y_2[\,.$$

We therefore have that

$$\forall y \in]h(\xi_1), y_2[, \quad y \text{ is neither a turning point nor a singular point.}$$

As h is a decreasing function on $]-\infty, -\sqrt{N}]$, all $\xi_1 \in]-\infty, -\sqrt{N}]$ except a countable number are continuity points. Choose then some $\tilde{\xi}_1$ to be a continuity point of h and $\tilde{x}_2^0 \in]h(\tilde{\xi}_1), y_2[$. By continuity of h, there exists a neighborhood $\tilde{V}$ of $(\tilde{x}_2^0, \tilde{\xi}_1)$ such that

$$\forall (x_2^0, \xi_1) \in \tilde{V}, x_2^0 - h(\xi_1) > 0\,.$$

The set $\{\big(x_1^0, \xi_1, x_2^0, (\varrho(x_2^0) - \xi_1^2)^{\frac{1}{2}}\big)\,;\ (x_1^0, \xi_1, x_2^0) \in \mathbf{R} \times \tilde{V}\big\}$ is a submanifold of $\mathbf{R} \times \mathbf{R}^* \times \mathbf{R}^2$ having codimension 1. Furthermore, for any initial data in this set, we have $\dot{x}_{2|t=0} > 0$ and a simple connexity argument shows that x_2^t is an increasing function of time. In particular $x_2^t \to y_2$ as $t \to \infty$. This proves Proposition 4.9 □

4.4. **Some examples in the betaplane approximation.** We are concerned here with the betaplane approximation, that is when $b(x_2) \equiv \beta x_2$. One has therefore

$$\widetilde{\tau}_R(\xi_1, x_2, \xi_2) = \bar{u}_1(x_2)\xi_1 + \frac{\beta\xi_1}{\xi_1^2 + \xi_2^2 + \beta^2 x_2^2}$$

and with the notation of Paragraph 4.2,

$$V_{\tau,\xi_1}(x_2) = \frac{\beta\xi_1}{\tau} - \xi_1^2 - \beta^2 x_2^2.$$

• In the absence of convection, one can characterize exactly the set Λ of initial data giving rise to trapped trajectories. One can notice that for τ such that $\tau\,\xi_1^{-1} \in]0, \beta\,\xi_1^{-2}]$, the energy surface $\mathcal{E}_{\tau,\xi_1}$ is simply the ellipse

$$\big\{\, (x_2, \xi_2) \,/\, \xi_2^2 + \beta^2\, x_2^2 = \beta\,\xi_1\,\tau^{-1} - \xi_1^2 \,\big\}\,.$$

Let us go through the previous analysis and study the trajectories in this situation. One notices that *fixed points* correspond to $x_2(t) = \xi_2(t) = 0$, with $x_1(t) = x_1^0$. There are no *asymptotic trajectories* (singular trajectories would correspond to $\tau = 0$, which is not possible here since $\xi_1 \neq 0$). Finally let us consider *periodic trajectories*. In order to get trapping one must check that if the energy level τ and the frequency ξ_1 are fixed, we have

$$\int_{x_{min}}^{x_{max}} \frac{\beta\left(\frac{\beta\,\xi_1}{\tau} - 2\,\xi_1^2\right)}{\sqrt{\frac{\beta\xi_1}{\tau} - \xi_1^2 - \beta^2\,x_2^2}}\, dx_2 = 0.$$

Let us choose for instance any $(x_1, \xi_1) \in \mathbf{R} \times \mathbf{R}^*$ and define the energy level $\tau = \beta/(2\xi_1)$. Then the integral is identically equal to zero hence the corresponding trajectory is trapped. It corresponds to $x_1^t = x_1^0$, and to x_2^t, ξ_2^t satisfying for all times ${\xi_2^t}^2 + \beta^2 {x_2^t}^2 = \xi_1^2$.

Such an explicit characterization cannot be obtained if there is some convection, in particular in physically relevant situations. Nevertheless, we are able to prove that, under suitable assumptions, Λ is not empty, and more precisely that it contains both singular trapped trajectories and periodic trapped trajectories.

• Proposition 4.9 shows that under the assumption that $\bar{u}_1 \in C_c^\infty(\mathbf{R})$ is a function which is not identically positive, with some zero of finite multiplicity, Λ contains a subset of codimension 1 of initial data giving rise to singular trapped trajectories.

• Finally let us construct periodic trajectories. We suppose to simplify the notation that $\beta = 1$, and that $\bar{u}_1$ has a local maximum at zero, with, say

$$0 < \bar{u}_1(0) < 2/3 \qquad \text{and} \qquad \bar{u}_1''(0) < -6.$$

Then we shall prove that the set Λ_{per} is nonempty and contains a submanifold of $\mathbf{R} \times \mathbf{R}^* \times \mathbf{R}^2$ which is of codimension 1. Let τ be an energy level. We notice that for η small enough and $y \in]-\eta, \eta[$

$$\bar{u}_1(y) < \bar{u}_1(0) < \frac{\tau}{\xi_1}$$

so there is no singular point for V_{τ,ξ_1} in the interval $]-\eta, \eta[$. Now let us show that there are two turning points inside $]-\eta, \eta[$ for η small enough; this will imply that there is a periodic trajectory of energy τ. Define the function

$$H_{\xi_1}(y) := \frac{\tau}{\xi_1} - \frac{1}{\xi_1^2 + y^2} - \bar{u}_1(y).$$

Then $H'_{\xi_1}(0) = 0$ and an easy computation shows that if η is chosen small enough, then for all $(y, \xi_1) \in]-\eta, \eta[\times [1, +\infty[$,

$$H''_{\xi_1}(y) = \frac{2\xi_1^2 - 8y^2}{(\xi_1^2 + y^2)^3} - \bar{u}_1''(y) \geq 1/2 > 0.$$

Now let us choose η such that $\sup\limits_{]-\eta,\eta[} \bar{u}_1''(y) < -3$, and let us define the number $\delta := 1 - \dfrac{\eta^2}{8}$.

We choose from now on $\xi_1 > 0$ and $\tau \equiv \tau(\xi_1) := \bar{u}_1(0)\xi_1 + \dfrac{\delta}{\xi_1}$ so that $H_{\xi_1}(0) = -\dfrac{1-\delta}{\xi_1^2} < 0$. The function H_{ξ_1} is decreasing on $]-\eta, 0[$ and increasing on the interval $]0, \eta[$; moreover

$$H_{\xi_1}(\pm\eta) \geq H_{\xi_1}(0) + \frac{\eta^2}{4} = -(1-\delta) + \frac{\eta^2}{4} \geq \frac{\eta^2}{8} > 0.$$

It follows that, for all $\xi_1 \in [1, +\infty[$, there are two points $x_{min}(\xi_1)$ (in $]-\eta, 0[$) and $x_{max}(\xi_1)$ (in $]0, \eta[$) satisfying the requirements of periodic trajectories, in the sense of Section 4.2.2.

According to the criterion (4.3), we define the following function on $[1, +\infty[$:

$$G(\xi_1) := \int_{x_{min}(\xi_1)}^{x_{max}(\xi_1)} \Big[\xi_1^2 - \frac{(\tau/\xi_1)}{2\big((\tau/\xi_1) - \bar{u}_1(y)\big)^2}\Big] V_{\tau,\xi_1}(y)^{-1/2}\, dy,$$

and let us prove it vanishes. We notice that recalling that $\bar{u}_1(0) < 2/3$,

$$\begin{aligned} G(1) &\geq \int_{x_{min}(1)}^{x_{max}(1)} \Big[1 - \Big(\bar{u}_1(0) + \delta\Big)\frac{1}{2\delta^2}\Big] V_{\tau,\xi_1}(y)^{-1/2}\, dy \\ &\geq \frac{6\delta^2 - 2 - 3\delta}{6\delta^2} \int_{x_{min}(1)}^{x_{max}(1)} V_{\tau,\xi_1}(y)^{-1/2}\, dy > 0 \end{aligned}$$

since $\eta \in]0,1[$. On the other hand, when ξ_1 goes to $+\infty$, we find that $x_{min}(\xi_1) \to 0$ and $x_{max}(\xi_1) \to 0$ so that recalling that $\tau(\xi_1) = \bar{u}_1(0)\xi_1 - \delta/\xi_1$,

$$G(\xi_1) \sim -\bar{u}_1(0)\, \frac{\xi_1^4}{2\delta^2} \int_{x_{min}(\xi_1)}^{x_{max}(\xi_1)} V_{\tau,\xi_1}(y)^{-1/2}\, dy,$$

so $\lim\limits_{\xi_1 \to +\infty} G(\xi_1) < 0$. By construction, the function G is smooth so there is some $\tilde{\xi}_1$ belonging to $]1, +\infty[$ such that $G(\tilde{\xi}_1) = 0$. We have generically (in τ) $G'(\tilde{\xi}_1) \neq 0$. This implies that $\Lambda_{per} \neq \emptyset$, where Λ_{per} is the subset of Λ giving rise to periodic trajectories, and that Λ_{per} contains a submanifold of $\mathbf{R} \times \mathbf{R}^* \times \mathbf{R}^2$ which is of codimension 1.

5. Study of the Poincaré waves

5.1. **The strategy.** In this section we want to prove the dispersion property of the Poincaré polarisation, namely result (2) of Theorem 1. We recall that the principal symbols $\tau_\pm$ are of order one, so one needs to study diffractive-type propagation, on a time scale of the order $1/\varepsilon^2$. Computing the bicharacterstics, as in the previous section, is therefore not enough to understand the classical flow. We shall instead rely on a spectral argument to prove that Poincaré waves do disperse, and escape from any compact set in the physical space. The first step consists in taking the Fourier transform in x_1 (recalling that the problem is invariant by translations in the x_1 direction). We also recall that the data is microlocalized on a compact set $\mathcal{C}$ such that ξ_1 is bounded away from zero. Since $\tau_\pm = \pm\sqrt{\xi_2^2 + b^2(x_2) + \xi_1^2}$, functional calculus implies that one can find classical pseudo-differential operators $H_\pm(\xi_1)$ of principal symbols $\xi_2^2 + b^2(x_2)$ such that $T_\pm = \pm\sqrt{H_\pm(\xi_1) + \xi_1^2}$. Let us now call $\lambda_\pm^k(\xi_1)$ and $\varphi_\pm^k(\xi_1; x_2)$ the eigenvalues and eigenfunctions of $H_\pm(\xi_1)$. The following proposition will be proved in the next paragraph.

Proposition 5.1. *Let ϕ be an eigenfunction of $H_\pm(\xi_1)$, microlocalized on an energy surface which interstects $\mathcal{C}$. Then ϕ and its associate eigenvalue λ are C^∞ functions of ξ_1. Moreover $\frac{1}{\varepsilon}\partial_{\xi_1}\lambda$ is bounded on compact sets in ξ_1.*

5.2. **Proof of Theorem 1(2) assuming Proposition 5.1.** Let us now carry out this program. We consider an initial data denoted φ^0, microlocalized in $\mathcal{C}$. One can take the Fourier transform in x_1 which gives

$$\varphi^0(x) = \frac{1}{\sqrt{2\pi\varepsilon}} \int \hat{\varphi}^0(\xi_1, x_2) e^{-i\frac{x_1\xi_1}{\varepsilon}}\, d\xi_1.$$

Now let us consider a coherent state (in Fourier variables) at (q,p) (see Appendix B), that is:

$$\varphi_{qp}(\xi_1) := \frac{1}{(\pi\varepsilon)^{\frac{1}{4}}} e^{i\frac{\xi_1 q}{\varepsilon}} e^{-\frac{(\xi_1 - p)^2}{2\varepsilon}}. \tag{5.1}$$

After decomposition onto coherent states we get

$$\varphi^0(x) = \frac{1}{\sqrt{2\pi\varepsilon}} \frac{1}{(\pi\varepsilon)^{\frac{1}{4}}} \int \tilde{\varphi}^0(q,p,x_2) e^{i\frac{\xi_1(q-x_1)}{\varepsilon}} e^{-\frac{(\xi_1-p)^2}{2\varepsilon}}\, dqdpd\xi_1$$

where $\tilde{\varphi}^0(q,p,x_2) := \left(\varphi_{qp}|\hat{\varphi}^0(\cdot,x_2)\right)_{L^2}$. We notice that the integral over p and q is, modulo $O(\varepsilon^\infty)$, on a compact domain due to the microlocalization assumption on φ^0. Finally decomposing onto the eigenfunctions $\varphi^k_\pm(\xi_1;x_2)$ gives

$$\varphi^0(x) = \frac{1}{\sqrt{2\pi\varepsilon}} \frac{1}{(\pi\varepsilon)^{\frac{1}{4}}} \sum_k \int \overline{\varphi}^0(q,p;k,\xi_1,x_2) \varphi^k_\pm(\xi_1;x_2) e^{i\frac{\xi_1(q-x_1)}{\varepsilon}} e^{-\frac{(\xi_1-p)^2}{2\varepsilon}}\, dqdpd\xi_1$$

where $\overline{\varphi}^0(q,p;k,\xi_1,x_2) := \left(\varphi^k_\pm(\xi_1;\cdot)|\tilde{\varphi}^0(q,p,\cdot)\right)_{L^2}$. Note that the dependence of $\overline{\varphi}^0$ on ξ_1 is only through the eigenfunction $\varphi^k_\pm$, so $\overline{\varphi}^0$ depends smoothly on ξ_1, as stated in Proposition 5.1.

The sum over k contains $O(\varepsilon^{-1})$ terms, due to the fact that $\lambda^k_\pm(\xi_1)$ remains in a finite interval (this will be made more precise in the next section, see Remark 5.3). Now it remains to propagate at time t/ε^2 this initial data, which gives rise to the following expression:

$$\frac{1}{\sqrt{2\pi\varepsilon}} \frac{1}{(\pi\varepsilon)^{\frac{1}{4}}} \sum_k \int \overline{\varphi}^0(q,p;k,\xi_1,x_2) \varphi^k_\pm(\xi_1;x_2) e^{i\frac{\xi_1(q-x_1)}{\varepsilon}} e^{-\frac{(\xi_1-p)^2}{2\varepsilon}} e^{\pm i(\lambda^k_\pm(\xi_1)+\xi_1^2)^{\frac{1}{2}}\frac{t}{\varepsilon^2}}\, dqdpd\xi_1$$

The stationary phase lemma then gives that this integral is $O(\varepsilon^\infty)$ except if there exists a stationary point, given by the conditions:

$$\xi_1 = p \text{ and } \varepsilon(x_1-q) \pm \frac{2\xi_1 + \partial_{\xi_1}\lambda^k_\pm(\xi_1)}{2\sqrt{\lambda^k_\pm(\xi_1)+\xi_1^2}} t = 0.$$

The second condition gives

$$2\sqrt{\lambda^k_\pm(\xi_1)+\xi_1^2}(x_1-q) = \mp\frac{1}{\varepsilon}\left(2p + \partial_{\xi_1}\lambda^k_\pm(\xi_1)\right) t.$$

Therefore, since $p \neq 0$ and the $\lambda^k_\pm$'s are bounded, with $\partial_{\xi_1}\lambda^k_\pm(\xi_1) = O(\varepsilon)$, there is no critical point for x_1 in a compact set. Proposition 5.1 therefore allows to apply the stationary phase lemma and to conclude the proof of result (2) of Theorem 1. Notice that the (fixed) losses in ε (namely the negative powers of ε appearing in the integrals and the number of k's in the sum) are compensated by the fact that the result is $O(\varepsilon^\infty)$; it is important at this point that as noticed above, the function $\overline{\varphi}^0$ depends smoothly on ξ_1.

5.3. **Proof of Proposition 5.1.** The first step of the proof consists in using the theory of normal forms in order to reduce the problem to the study of functions of the harmonic oscillator (as in Bohr-Sommerfeld quantization, paying special attention to the dependence on ξ_1). The second step then consists in checking that the eigenvectors and eigenvalues have the required dependence on ξ_1. In the following we shall only deal with H_+ to simplify, and we shall write $H := H_+$. The first step relies on the following lemma.

Lemma 5.2. *Suppose that ξ_1 lies in a compact set away from zero. There is an elliptic Fourier Integral Operator $U : \mathcal{D}(\mathbf{R}) \to \mathcal{D}(\mathbf{R})$, independent of ξ_1, and a pseudodifferential operator $V(\xi_1)$ with C^∞ symbol in ξ_1 such that microlocally in any compact set $K \subset T^*\mathbf{R}$ one has $(VU)^* = (VU)^{-1}$ and*

$$VUH(VU)^{-1} = f(-\varepsilon^2\partial_2^2 + x_2^2;\xi_1;\varepsilon) + O(\varepsilon^\infty)$$

where

$$f(I;\xi_1;\varepsilon) = f_0(I) + \sum \varepsilon^j f_j(I;\xi_1;\varepsilon)$$

while f_0 *is a smooth bijection on* $\mathbf{R}^+$ *and the* f_j*'s are* C^∞ *functions of* I *and* ξ_1.

Proof. The proof consists in using techniques linked to the isochore Morse lemma (see [6], [30]). Actually we introduce a canonical change of variables (the corresponding operator being the FIO U) allowing to pass from the variables (x_2, ξ_2) to action-angle variables (see [1] for instance). Let us make this first step more precise : we recall that the action variables are given by

$$I := \oint \xi_2 \, dx_2,$$

where the integral is taken on a constant energy curve $H = h$. The angle variables are then given by solving $\theta = \partial_I S(I, \xi_2)$, where K is the hamiltonian in the new variables, which only depends on I, and where S is defined by

$$dS_{I=constant} = \xi_2 dx_2.$$

Note that this is a *global* change of variables. It is now well known (see [7] for instance) that such a canonical change of variables is associated with an FIO U (independent of ξ_1 since the principal symbol of H does not depend on ξ_1) such that

$$UHU^{-1} = f^0(-\varepsilon^2\partial_2^2 + x_2^2) + \varepsilon F_1 \tag{5.2}$$

where f^0 is a smooth, global bijection on $\mathbf{R}^+$. In this case we can actually write (see [7]) the following formula for U: for any L^2 function φ

$$U\varphi(x_2) = \frac{1}{(2\pi\varepsilon)^{\frac{3}{2}}} \int e^{i\frac{S(x_2,\xi_2)-y\xi_2}{\varepsilon}} a_0(x_2, \xi_2, \varepsilon) u(y) \, dy d\xi_2$$

where a_0 is constructed so that U is unitary (up to $O(\varepsilon)$). Using those new coordinates and the formula giving U as well as the formula for the principal symbol of the adjoint (which is here the inverse) given in Appendix B, it is not difficult to show that (5.2) holds. Once the function f^0 is obtained we proceed by induction: we first look for a symbol q_1 such that $Q_1 := \mathrm{Op}_\varepsilon^0(q_1)$ satisfies

$$e^{iQ_1} UHU^{-1} e^{-iQ_1} = f^{0,1}(-\varepsilon^2\partial_2^2 + x_2^2) + \varepsilon^2 F_2.$$

In order to compute q_1 we notice that

$$e^{iQ_1} UHU^{-1} e^{-iQ_1} = UHU^{-1} + \int_0^t \partial_t H_t \, dt$$

where $H_t := e^{itQ_1} UHU^{-1} e^{-itQ_1}$. One sees easily that

$$\begin{aligned} \partial_t H_t &= i[Q_1, H_t] \\ &= e^{itQ_1}[Q_1, f^0(-\varepsilon^2\partial_2^2 + x_2^2) + \varepsilon F_1] e^{-itQ_1} \end{aligned}$$

so the principal symbol of $\partial_t H_t$ is therefore the Poisson bracket $\varepsilon\{q_1, f^0(x_2^2 + \xi_2^2)\}$. One then remarks that the equation

$$\{q, f^0(x_2^2 + \xi_2^2)\} = g$$

has a solution if and only if g has zero mean value (in action-angle variables): one has indeed necessarily

$$\partial_\theta q = \frac{Ig}{\partial_I f}.$$

Note that if f and g are smooth, then so is q since $\partial_I f > 0$ everywhere. It follows that one can find q_1 simply by solving

$$\{q_1, f^0(x_2^2+\xi_2^2)\} = f_1 - \overline{f}_1$$

where $\overline{f}_1$ is the average of f_1

$$\overline{f}_1(I) = \frac{1}{2\pi}\int_0^{2\pi} f_1(R_\theta(x_2,\xi_2);\xi_1)\,d\theta$$

where R_θ denotes the rotation of angle θ. This implies that with this choice of q_1 and writing $Q_1 := \mathrm{Op}_\varepsilon^0(q_1)$ one has

$$e^{iQ_1}UHU^{-1}e^{-iQ_1} = (f^0+\varepsilon\overline{f}_1)(-\varepsilon^2\partial_2^2+x_2^2)+\varepsilon^2F_2.$$

One proceeds similary at all orders. □

The next step consists in using that lemma to check that the eigenvalues and eigenfunctions enjoy the expected smoothness properties. Actually this is rather straightforward since if (ψ,λ) satisfy

$$H\psi = \lambda\psi + O(\varepsilon^\infty)$$

then $\widetilde{\psi} := VU\psi$ satisfies

$$VUH(VU)^{-1}\widetilde{\psi} = \lambda\widetilde{\psi} + O(\varepsilon^\infty)$$

hence

$$f(-\varepsilon^2\partial_2^2+x_2^2)\widetilde{\psi} = \lambda\widetilde{\psi} + O(\varepsilon^\infty).$$

This implies that

$$\lambda = \lambda_n = f\left(\varepsilon\big(n+\frac{1}{2}\big);\xi_1;\varepsilon\right) \quad \text{and} \quad \widetilde{\psi} = \widetilde{\psi}_n = \frac{1}{\varepsilon^{\frac{1}{4}}}h_n\left(\frac{x_2}{\sqrt{\varepsilon}}\right)$$

where h_n is the n-th Hermite function. The conclusion follows recalling that $f(I;\xi_1;\varepsilon) = f_0(I) + \sum \varepsilon^j f_j(I;\xi_1;\varepsilon)$ (where f_0 does not depend on ξ_1) and that each eigenfunction is microlocalized in a compact set.

Remark 5.3. *Since we have a complete spectral description of $T_\pm$, with discrete spectrum for each given ξ_1, it is obvious that if the initial data is microlocalized in (ξ_1,x_2,ξ_2), then the solution to the equation $\varepsilon^2\partial_t\varphi = iT_\pm\varphi$ remains microlocalized in the set of energy surfaces containing (ξ_1,x_2,ξ_2). Notice also that there are $O(1/\varepsilon)$ eigenvalues in a compact energy surface.*

6. Diagonalization

In this section we shall prove that the scalar propagators defined in Section 3 correspond indeed to a diagonalization of the original linear system (2.1). Let us prove the following proposition.

Proposition 6.1. *Consider a compact set $\mathcal{K} \subset \mathbf{R}^* \times T^*\mathbf{R}$. With the notation of Definitions 3.3 and 3.6, the operator $\Pi := (\Pi_- \,\Pi_R\, \Pi_+)$ maps continuously H_ε^∞ functions, microlocalized in $\mathbf{R}\times\mathcal{K}$ and satisfying (2.3), onto H_ε^∞. Moreover it is left-invertible modulo ε^∞, and its left inverse Q (modulo ε^∞) maps continuously those functions onto H_ε^∞.*

Remark 6.2. *This proposition, along with Proposition 4.6 and Remark 5.3 showing the propagation of the microlocal support (in (ξ_1, x_2, ξ_2)) of Rossby and Poincaré modes, prove the first part of Theorem 1.*

Proof. The main step consists in showing that Π does have a left inverse, and in computing its principal symbol. The construction of a left inverse can be done symbolically as follows. We first compute the matrix-principal symbol $\mathcal{P}_0$ of Π. One gets

$$\mathcal{P}_0 = \begin{pmatrix} \dfrac{-\xi_1\sqrt{\xi^2+b^2} - i\xi_2 b}{\xi_1^2+b^2} & -\dfrac{ib}{\xi_1} & \dfrac{\xi_1\sqrt{\xi^2+b^2} - i\xi_2 b}{\xi_1^2+b^2} \\ 1 & -\dfrac{\xi_2}{\xi_1} & 1 \\ \dfrac{\xi_1\xi_2 + ib\sqrt{\xi^2+b^2}}{\xi_1^2+b^2} & 1 & \dfrac{\xi_1\xi_2 - ib\sqrt{\xi^2+b^2}}{\xi_1^2+b^2} \end{pmatrix}.$$

This shows that Π maps microlocalized functions in $\mathbf{R}\times\mathcal{K}$ onto H^∞_ε. A simple computation shows that

$$|\det \mathcal{P}_0| = \frac{2(\xi^2+b^2)^{\frac{3}{2}}}{(\xi_1^2+b^2)|\xi_1|} \geq 2,$$

therefore $\mathcal{Q}_0 = \mathcal{P}_0^{-1}$ exists. Let us call Q_0 the matrix obtained by Weyl quantization (term by term) of $\mathcal{Q}_0$. By symbolic calculus we have that:

$$Q_0\Pi = \mathrm{Id} + \varepsilon I_1$$

Let us call $\mathcal{Q}_1 = -\mathcal{I}_1\mathcal{Q}_0^{-1}$, where $\mathcal{I}_1$ is the (matrix) principal symbol of I_1. Again by symbolic calculus we have that

$$(Q_0 + \varepsilon Q_1)\,\Pi = \mathrm{Id} + \varepsilon^2 I_2$$

where Q_1 has principal symbol $\mathcal{Q}_1$. Defining now $\mathcal{Q}_2 = -\mathcal{I}_2\mathcal{Q}_0^{-1}$ we get

$$\left(Q_0 + \varepsilon Q_1 + \varepsilon^2 Q_2\right)\Pi = \mathrm{Id} + \varepsilon^3 I_3$$

where Q_2 has principal symbol $\mathcal{Q}_2$, and so on. This allows to invert the matrix Π up to $O(\varepsilon^\infty)$, and the principal symbol of the (approximate) inverse matrix Q is given by $\mathcal{Q}_0$: we have

$$\mathcal{Q}_0 = \frac{1}{2(\xi^2+b^2)}\begin{pmatrix} ib\xi_2 - \xi_1\sqrt{\xi^2+b^2} & \xi_1^2+b^2 & -ib\sqrt{\xi^2+b^2} + \xi_1\xi_2 \\ 2ib\xi_1 & -2\xi_1\xi_2 & 2\xi_1^2 \\ ib\xi_2 + \xi_1\sqrt{\xi^2+b^2} & \xi_1^2+b^2 & ib\sqrt{\xi^2+b^2} + \xi_1\xi_2 \end{pmatrix}.$$

Note that each term of the expansion of the symbol of Q is a polynomial of increasing order in ξ, but that is not a problem due to the microlocalization assumption. The operator Q therefore clearly maps continuously H^∞_ε onto itself. The proposition is proved. □

7. Control of the nonlinear terms for a weak coupling

We are now interested in describing the behaviour of our initial nonlinear system, which includes the effect of the convection by the unknown (u_1, u_2). We shall prove Theorem 2 in this section. We recall that the system reads as follows:

$$\varepsilon^2\partial_t U + A(x,\varepsilon D_x)U + \varepsilon^{3+\eta}S_1(U)\varepsilon\partial_1 U + \varepsilon^{3+\eta}S_2(U)\varepsilon\partial_2 U = 0, \quad \eta \geq 0 \tag{7.1}$$

with

$$S_1(U) = \begin{pmatrix} u_1 & \frac{1}{2}u_0 & 0 \\ \frac{1}{2}u_0 & u_1 & 0 \\ 0 & 0 & u_1 \end{pmatrix}, \qquad S_2(U) = \begin{pmatrix} u_2 & 0 & \frac{1}{2}u_0 \\ 0 & u_2 & 0 \\ \frac{1}{2}u_0 & 0 & u_2 \end{pmatrix} \tag{7.2}$$

and $U = (u_0, u_1, u_2)$. The usual theory of symmetric hyperbolic systems provides the local existence of a solution to (7.1) in $H^s(\mathbf{R}^2)$ for $s > 2$, on a time interval depending on ε a priori. Because of the semiclassical framework (fixed by the form of the initial data), it is actually natural to rather consider ε-derivatives. Moreover, as the derivative with respect to x_2 does not commute with the singular perturbation, we expect even the semiclassical Sobolev norms to grow like $\exp\left(\frac{Ct}{\varepsilon}\right)$, and therefore the life span of the solutions to (7.1) to be non uniform with respect to ε. That is the reason why the W^s_ε spaces were introduced in (1.7).

7.1. **Propagation of regularity for the linear singular perturbation problem.** Let us first remark that derivatives with respect to x_1 do commute with $A(x, \varepsilon D_x)$, so we can propagate as much regularity in x_1 as needed. Extending a result by Dutrifoy, Majda and Schochet [10] obtained in the particular case when $b(x_2) = \beta x_2$, we will actually prove that there is an operator of principal symbol $(\xi_2^2 + b^2)\mathrm{Id}$ which "almost commutes" with $A(x, \varepsilon D_x)$ in the semiclassical regime.

• The first step, as in [10], is to perform the following orthogonal change of variable

$$\tilde{U} := \left(\frac{u_0 + u_1}{\sqrt{2}}, \frac{u_0 - u_1}{\sqrt{2}}, u_2\right)$$

in order to produce the generalized creation and annihilation operators

$$L_\pm := \frac{1}{\sqrt{2}}\big(\varepsilon\partial_2 \mp b\big)\,.$$

The system (7.1) can indeed be rewritten

$$\varepsilon^2\partial_t\tilde{U} + \tilde{A}(x, \varepsilon D_x)\tilde{U} + \varepsilon^3\tilde{S}_1(\tilde{U})\varepsilon\partial_1\tilde{U} + \varepsilon^3\tilde{S}_2(\tilde{U})\varepsilon\partial_2\tilde{U} = 0$$

with

$$\tilde{A}(x, \varepsilon D_x) := \begin{pmatrix} \varepsilon\bar{u}_1\varepsilon\partial_1 + \varepsilon\partial_1 & 0 & L_+ + \frac{\varepsilon^2}{\sqrt{2}}\bar{u}_1' \\ 0 & \varepsilon\bar{u}_1\varepsilon\partial_1 - \varepsilon\partial_1 & L_- - \frac{\varepsilon^2}{\sqrt{2}}\bar{u}_1' \\ L_- & L_+ & \varepsilon\bar{u}_1\varepsilon\partial_1 \end{pmatrix},$$

and

$$\tilde{S}_1(\tilde{U}) := \begin{pmatrix} \dfrac{3\tilde{u}_0 - \tilde{u}_1}{2\sqrt{2}} & 0 & 0 \\ 0 & \dfrac{\tilde{u}_0 - 3\tilde{u}_1}{2\sqrt{2}} & 0 \\ 0 & 0 & \dfrac{\tilde{u}_0 - \tilde{u}_1}{\sqrt{2}} \end{pmatrix}, \qquad \tilde{S}_2(\tilde{U}) := \begin{pmatrix} \tilde{u}_2 & 0 & \dfrac{\tilde{u}_0 + \tilde{u}_1}{4} \\ 0 & \tilde{u}_2 & \dfrac{\tilde{u}_0 + \tilde{u}_1}{4} \\ \dfrac{\tilde{u}_0 + \tilde{u}_1}{4} & \dfrac{\tilde{u}_0 + \tilde{u}_1}{4} & \tilde{u}_2 \end{pmatrix}.$$

• Next, remarking that $[\varepsilon^2\partial_2^2 - b^2, \varepsilon\partial_2 \pm b] = \pm 2\varepsilon b'(\varepsilon\partial_2 \pm b) \pm \varepsilon^2 b''$, we introduce the operator

$$D_\varepsilon := \begin{pmatrix} \varepsilon^2\partial_2^2 - b^2 + 2\varepsilon b' & 0 & 0 \\ 0 & \varepsilon^2\partial_2^2 - b^2 - 2\varepsilon b' & 0 \\ 0 & 0 & \varepsilon^2\partial_2^2 - b^2 \end{pmatrix}.$$

We notice that D_ε is a scalar operator at leading order. Moreover one can compute the commutator $[D_\varepsilon, \tilde{A}(x, \varepsilon D_x)]$: we find

$$[D_\varepsilon, \tilde{A}(x, \varepsilon D_x)] = \begin{pmatrix} [\varepsilon^2\partial_2^2, \varepsilon\bar{u}_1]\varepsilon\partial_1 & 0 & \frac{\varepsilon^2}{\sqrt{2}}\big([\varepsilon^2\partial_2^2, \bar{u}_1'] + 2\varepsilon b'\bar{u}_1' - b''\big) \\ 0 & [\varepsilon^2\partial_2^2, \varepsilon\bar{u}_1]\varepsilon\partial_1 & -\frac{\varepsilon^2}{\sqrt{2}}\big([\varepsilon^2\partial_2^2, \bar{u}_1'] - 2\varepsilon b'\bar{u}_1' - b''\big) \\ -\frac{\varepsilon^2}{\sqrt{2}}b'' & \frac{\varepsilon^2}{\sqrt{2}}b'' & [\varepsilon^2\partial_2^2, \varepsilon\bar{u}_1]\varepsilon\partial_1 \end{pmatrix}$$

• At this stage, we have proved that

$$[D_\varepsilon, \tilde{A}(x, \varepsilon D_x)] = O(\varepsilon^2(\mathrm{Id} - \varepsilon^2\partial_1^2 - D_\varepsilon)) \tag{7.3}$$

meaning that the commutator $[D_\varepsilon, \tilde{A}(x, \varepsilon D_x)]$ is of order $O(\varepsilon^2)$ with respect to the elliptic operator $\mathrm{Id} - \varepsilon^2\partial_1^2 - D_\varepsilon$. That implies that the regularity of the solution to the linear equation

$$\varepsilon^2\partial_t V + \tilde{A}(x, \varepsilon D_x)V = 0$$

can be controlled by an application of Gronwall's lemma: one has

$$\varepsilon^2\|(Id - \varepsilon^2\partial_1^2 - D_\varepsilon)V(t)\|_{L^2}^2 \leq \varepsilon^2\|(Id - \varepsilon^2\partial_1^2 - D_\varepsilon)V_0\|_{L^2}^2 + C\varepsilon^2\int_0^t \|(Id - \varepsilon^2\partial_1^2 - D_\varepsilon)V(s)\|_{L^2}^2 ds\,,$$

where C depends on the $W^{2,\infty}$ norms of $\bar{u}_1$ and b, so

$$\|(Id - \varepsilon^2\partial_1^2 - D_\varepsilon)V(t)\|_{L^2}^2 \leq C\|(Id - \varepsilon^2\partial_1^2 - D_\varepsilon)V_0\|_{L^2}^2\, e^{Ct}.$$

7.2. **Uniform a priori estimates for the nonlinear equation.** Since the extended harmonic oscillator controls two derivatives in x_2, we get a control on the Lipschitz norm of U of the type

$$\|\varepsilon\partial_j U\|_{L^\infty} \leq \frac{C}{\varepsilon}(\|D_\varepsilon^2 U\|_{L^2} + \|\varepsilon^4\partial_1^4 U\|_{L^2} + \|U\|_{L^2})\,. \tag{7.4}$$

As D_ε is a scalar differential operator at leading order in ε, the antisymmetry of the higher order nonlinear term is preserved. More precisely, we have, using the Leibniz formula,

$$\begin{aligned}\varepsilon^2\partial_t D_\varepsilon\tilde{U} + \tilde{A}(x, \varepsilon D_x)D_\varepsilon\tilde{U} + \varepsilon^3\tilde{S}_1(\tilde{U})\varepsilon\partial_1 D_\varepsilon\tilde{U} + \varepsilon^3\tilde{S}_2(\tilde{U})\varepsilon\partial_2 D_\varepsilon\tilde{U} \\ = -[D_\varepsilon, \tilde{A}(x, \varepsilon D_x)]\tilde{U} - \varepsilon^3\tilde{S}_2(\tilde{U})[D_\varepsilon, \varepsilon\partial_2]\tilde{U} - \varepsilon^3[D_\varepsilon, \tilde{S}_j(\tilde{U})]\varepsilon\partial_j\tilde{U}\end{aligned}$$

as well as

$$\begin{aligned}\varepsilon^2\partial_t D_\varepsilon^2\tilde{U} + \tilde{A}(x, \varepsilon D_x)D_\varepsilon^2\tilde{U} + \varepsilon^3\tilde{S}_1(\tilde{U})\varepsilon\partial_1 D_\varepsilon^2\tilde{U} + \varepsilon^3\tilde{S}_2(\tilde{U})\varepsilon\partial_2 D_\varepsilon^2\tilde{U} = -[D_\varepsilon^2, \tilde{A}(x, \varepsilon D_x)]\tilde{U} \\ -\varepsilon^3\tilde{S}_2(\tilde{U})[D_\varepsilon, \varepsilon\partial_2]D_\varepsilon\tilde{U} - \varepsilon^3[D_\varepsilon, \tilde{S}_j(\tilde{U})]\varepsilon\partial_j D_\varepsilon\tilde{U} \\ +D_\varepsilon\left(-\varepsilon^3\tilde{S}_2(\tilde{U})[D_\varepsilon, \varepsilon\partial_2]\tilde{U} - \varepsilon^3[D_\varepsilon, \tilde{S}_j(\tilde{U})]\varepsilon\partial_j\tilde{U}\right).\end{aligned}$$

and in the same way, for $1 \leq \ell \leq 4$,

$$\begin{aligned}\varepsilon^2\partial_t(\varepsilon\partial_1)^\ell\tilde{U} + \tilde{A}(x, \varepsilon D_x)(\varepsilon\partial_1)^\ell\tilde{U} + \varepsilon^3\tilde{S}_1(\tilde{U})(\varepsilon\partial_1)^{\ell+1}\tilde{U} + \varepsilon^3\tilde{S}_2(\tilde{U})\varepsilon\partial_2(\varepsilon\partial_1)^\ell\tilde{U} \\ = -\varepsilon^4\sum_{k=1}^{\ell} C_4^\ell(\varepsilon\partial_1)^\ell\tilde{S}_j(\tilde{U})\varepsilon\partial_j(\varepsilon\partial_1)^{\ell-k}\tilde{U}\,.\end{aligned}$$

In all cases, the terms of higher order disappear by integration in x and the other terms are controlled with the following trilinear estimate (writing generically $\tilde{Q}(\tilde{U})$ for all the nonlinearities): for all $0 \leq k \leq 2$ and all $0 \leq \ell \leq 4$

$$\begin{aligned} &| < D_\varepsilon^k \tilde{U} | D_\varepsilon^k \tilde{Q}(\tilde{U}) > | + | < (\varepsilon\partial_1)^\ell \tilde{U} | (\varepsilon\partial_1)^\ell \tilde{Q}(\tilde{U}) > | \\ &\qquad \leq C \|\tilde{U}\|_{W^{1,\infty}_\varepsilon} (\|D_\varepsilon^2 \tilde{U}\|_{L^2} + \|(\varepsilon\partial_1)^4 \tilde{U}\|_{L^2} + \|\tilde{U}\|_{L^2})^2 \\ &\qquad \leq \frac{C}{\varepsilon} \Big(\|D_\varepsilon^2 \tilde{U}\|_{L^2} + \|(\varepsilon\partial_1)^4 \tilde{U}\|_{L^2} + \|\tilde{U}\|_{L^2} \Big)^3. \end{aligned} \tag{7.5}$$

Remark 7.1. *Note that because of the bad embedding inequality* $\|\nabla U\|_{L^\infty} \leq \frac{1}{\varepsilon}\|U\|_{W^4_\varepsilon}$, *we lose one power of* ε, *which seems not to be optimal considering for instance the fast oscillating functions* $x_2 \mapsto \exp\left(\frac{ik_2x_2}{\varepsilon}\right)$. *A challenging question in order to apply semiclassical methods to nonlinear problems is to determine appropriate functional spaces (in the spririt of* [23]*) which measures on the one hand the Sobolev regularity of the amplitudes, and on the other hand the oscillation frequency.*

We are finally able **to obtain a uniform life span** for the weakly nonlinear system, thus proving result (1a) of Theorem 2. Indeed combining the trilinear estimate (7.5) and the commutator estimate (7.3), we obtain the following Gronwall inequality

$$\varepsilon^2 \frac{d}{dt} \Big(\|D_\varepsilon^2 \tilde{U}\|_{L^2}^2 + \|(\varepsilon\partial_1)^4 \tilde{U}\|_{L^2}^2 + \|\tilde{U}\|_{L^2}^2 \Big) \leq C\varepsilon^2 \Big(1 + \|D_\varepsilon^2 \tilde{U}\|_{L^2} + \|(\varepsilon\partial_1)^4 \tilde{U}\|_{L^2} + \|\tilde{U}\|_{L^2} \Big)^3$$

from which we deduce the uniform a priori estimate

$$\|D_\varepsilon^2 \tilde{U}\|_{L^2}^2 + \|(\varepsilon\partial_1)^4 \tilde{U}\|_{L^2}^2 + \|\tilde{U}\|_{L^2}^2 \leq (C_0 - Ct)^{-2}$$

where C_0 depends only on the initial data. Such an estimate shows that the life span of the solutions to (7.1) is at least $T^* = C_0/C$.

7.3. **Approximation by the linear dynamics.** In this paragraph we shall prove results (1b) and (2) of Theorem 2. The proof of both results relies on standard energy estimates. We have

$$\varepsilon^2 \partial_t (U_\varepsilon - V_\varepsilon) + A(x, \varepsilon D_x)(U_\varepsilon - V_\varepsilon) + \varepsilon^{3+\eta} \tilde{S}_1(U_\varepsilon) \varepsilon\partial_1 U_\varepsilon + \varepsilon^{3+\eta} \tilde{S}_2(U_\varepsilon) \varepsilon\partial_2 U_\varepsilon = 0$$

• If $\eta = 0$ and $\varepsilon V_\varepsilon \to 0$ in L^∞, we use the decomposition

$$\varepsilon^2 \partial_t (U_\varepsilon - V_\varepsilon) + A(x, \varepsilon D_x)(U_\varepsilon - V_\varepsilon) + \varepsilon^3 (\tilde{S}_j(U_\varepsilon) - \tilde{S}_j(V_\varepsilon)) \varepsilon\partial_j U_\varepsilon + \varepsilon^3 \tilde{S}_j(V_\varepsilon) \varepsilon\partial_j U_\varepsilon = 0$$

and obtain the following L^2 estimate

$$\begin{aligned} \frac{\varepsilon^2}{2} \frac{d}{dt} \|U_\varepsilon - V_\varepsilon\|_{L^2}^2 &\leq 3\varepsilon^3 \|\varepsilon\partial_j U_\varepsilon\|_{L^\infty} \|U_\varepsilon - V_\varepsilon\|_{L^2}^2 + 3\varepsilon^3 \|V_\varepsilon\|_{L^\infty} \|\varepsilon\partial_j U_\varepsilon\|_{L^2} \|U_\varepsilon - V_\varepsilon\|_{L^2} \\ &\leq C\varepsilon^2 (\varepsilon \|\varepsilon\partial_j U_\varepsilon\|_{L^\infty} + \|\varepsilon\partial_j U_\varepsilon\|_{L^2}^2) \|U_\varepsilon - V_\varepsilon\|_{L^2}^2 + C\varepsilon^2 (\varepsilon \|V_\varepsilon\|_{L^\infty})^2 \end{aligned}$$

from which we conclude by Gronwall's lemma

$$\|U_\varepsilon - V_\varepsilon\|_{L^2}^2 \leq C \int_0^t (\varepsilon \|V_\varepsilon(s)\|_{L^\infty})^2 \exp C \left(\int_s^t (\varepsilon \|\varepsilon\partial_j U_\varepsilon\|_{L^\infty} + \|\varepsilon\partial_j U_\varepsilon\|_{L^2}^2) d\sigma \right) ds$$

on $[0, T^*[$, and that proves result (1b).

• If $\eta > 0$, the same arguments show that the life span of the solutions to (2.5) tends to infinity as $\varepsilon \to 0$: $T_\varepsilon \geq C\varepsilon^{-\eta}$, and that these solutions are uniformly bounded in W^4_ε on any finite time interval. Furthermore, on any finite time interval $[0, T]$, the previous energy estimate gives

$$\frac{\varepsilon^2}{2}\frac{d}{dt}\|U_\varepsilon - V_\varepsilon\|^2_{L^2} \leq C\varepsilon^{3+\eta}\|\varepsilon\partial_j U_\varepsilon\|_{L^\infty}\|U_\varepsilon\|^2_{L^2}\,,$$

from which we deduce

$$\|U_\varepsilon - V_\varepsilon\|^2_{L^2} \leq C\varepsilon^\eta \int_0^t \varepsilon\|\varepsilon\partial_j U_\varepsilon(s)\|_{L^\infty}\|U_\varepsilon(s)\|^2_{L^2}ds\,.$$

Result (2) of Theorem 2 is proved.

Appendix A. A diagonalization theorem

In this appendix we shall state and prove the crucial theorem allowing to diagonalize semi-classically the matrix of pseudo-differential operators $A(x_2, \varepsilon D)$. The construction of Rossby and Poincaré modes (see Propositions 3.2 and 3.5 in Section 3) are direct corollaries of that theorem. We refer to Appendix B for the notation and results of semi-classical analysis used in this paragraph. We consider an order function g on $\mathbf{R}^{2d}$, and a symbol $h(X;\varepsilon,\tau)$ in $S_{2d}(g)$, depending polynomially on ε and τ, where we have defined $X = (x,\xi) \in \mathbf{R}^{2d}$. The statement is the following.

Theorem 3. *Let g be an order function on $\mathbf{R}^{2d}$, and let $h(X;\varepsilon,\tau)$ be a classical symbol in $S_{2d}(g)$, depending polynomially on ε and τ: there is an integer N_0 such that*

$$\forall X \in \mathbf{R}^{2d}, \quad h(X;\varepsilon,\tau) = \sum_{j,\ell=0}^{N_0} h^{j,\ell}(X)\varepsilon^\ell\tau^j, \tag{A.1}$$

where the symbols $h^{j,\ell}(X)$ belong to $S_{2d}(g)$. Let $\tilde\tau_\varepsilon = \tilde\tau_\varepsilon(X)$ be a root of the polynomial $h(X;\varepsilon,\cdot)$ which can be written for some $\nu \in \mathbf{N}$, $\tilde\tau_\varepsilon(X) = \sum_{k=0}^{\infty}\varepsilon^{\nu+k}\tau_k(X) + O(\varepsilon^\infty)$, with $\tau_0 \neq 0$ and where τ_0 is a symbol. Finally let $h_0(X,\tau)$ be the principal symbol of h, satisfying the following assumption:

$$\exists C > 0,\ \forall X \in \mathbf{R}^{2d},\ \forall\varepsilon \in]0,1[, |\partial_\tau h_0(X;\varepsilon^\nu\tau_0)| \geq C. \tag{A.2}$$

Let $\mathcal{K}$ be a compact subset of $\mathbf{R}^{2d}$. Then there is a pseudo-differential operator T of principal symbol $\varepsilon^\nu\tau_0$ such that if $\psi \in \mathcal{S}'(\mathbf{R}^d)$ is microlocalized in $\mathcal{K}$ and satisfies $i\partial_s\psi = T\psi$ then

$$\mathrm{Op}^0_\varepsilon(h(\varepsilon, i\partial_s))\psi = O(\varepsilon^\infty) \quad in \quad L^2(\mathbf{R}^d). \tag{A.3}$$

Proof. The idea of the proof of the theorem is the following. Let us define a smooth function χ, compactly supported in $\mathbf{R}^{2d}$, identically equal to one on $\mathcal{K}$. Then for any integer N, we shall compute recursively the coefficients of the symbol $\tau^N_\varepsilon := \varepsilon^\nu\tau_0 + \sum_{k=1}^N \varepsilon^{\nu+k}\tau^-_k$ so $T_\chi := \mathrm{Op}^1_\varepsilon(\tau^\chi_\varepsilon)$ satisfies the required property, where τ^χ_ε, unique up to $O(\varepsilon^\infty)$, is given by

$$\tau^\chi_\varepsilon := \varepsilon^\nu\tau_0 + \sum_{k=1}^{\infty}\varepsilon^{\nu+k}\chi\tau^-_k + O(\varepsilon^\infty).$$

This allows to replace the root $\widetilde{\tau}_\varepsilon$ by an actual symbol. The above strategy will be achieved in the following way. We notice that if $i\partial_s\psi = T_\chi\psi$ then of course

$$i\partial_s\psi = \mathrm{Op}^1_\varepsilon(\tau_\varepsilon^{\chi,N})\psi + O(\varepsilon^{N+\nu+1}), \quad \text{where} \quad \tau_\varepsilon^{\chi,N} := \varepsilon^\nu\tau_0 + \sum_{k=1}^N \varepsilon^{\nu+k}\chi\tau_k^-.$$

If moreover ψ is microlocalized in $\mathcal{K}$, then by definition of χ one has using (B.3-B.4) and the fact that χ is identically equal to one over $\mathcal{K}$ $i\partial_s\psi = \mathrm{Op}^1_\varepsilon(\tau_\varepsilon^N)\psi + O(\varepsilon^{N+\nu+1})$. This means one can (and shall) compute recursively $(\tau_k^-)_{1\le k\le N}$ so that

$$\mathrm{Op}^0_\varepsilon(h(\varepsilon, i\partial_s))\psi = O(\varepsilon^{N+1}), \quad \text{when} \quad i\partial_s\psi = \mathrm{Op}^1_\varepsilon(\tau_\varepsilon^N)\psi + O(\varepsilon^{N+1}).$$

Note that it is convenient in the computations to compute T_χ as the "right"-quantization of the symbol τ_ε^χ. Now let us carry out the algebraic computations allowing to achieve the result. We shall start by dealing with the case when $\nu \neq 0$, as the computations can be carried out in an easier way, and then we shall discuss the case when $\nu = 0$. Recalling that $h(X;\varepsilon,\tilde{\tau}_\varepsilon(X)) = 0$ we infer that $h^{0,\ell} \equiv 0$ if $\ell < \nu$ and

$$h^{0,\nu}(x,\xi) + h^{10}(x,\xi)\tau_0(x,\xi) = 0. \tag{A.4}$$

Now we recall that for any ψ in $\mathcal{S}'(\mathbf{R}^d)$, one has

$$\mathrm{Op}^0_\varepsilon(h(\varepsilon, i\partial_s))\psi(x) = (2\pi\varepsilon)^{-d}\int_{\mathbf{R}^{2d}} e^{i\frac{(x-y)\cdot\xi}{\varepsilon}} h(x,\xi;\varepsilon,i\partial_s)\psi(y)\,dyd\xi.$$

The above integral, as all the ones appearing in this proof, is to be understood in the distributional sense. So if τ_ε and ψ are such that $\mathrm{Op}^1_\varepsilon(\tau_\varepsilon^\chi)\psi = i\partial_s\psi$ with ψ microlocalized in $\mathcal{K}$,

$$\begin{aligned}
\mathrm{Op}^0_\varepsilon(h(\varepsilon, i\partial_s))\psi(x) &= (2\pi\varepsilon)^{-d}\int_{\mathbf{R}^{2d}} e^{i\frac{(x-y)\cdot\xi}{\varepsilon}} h(x,\xi;\varepsilon,\mathrm{Op}^1_\varepsilon(\tau_\varepsilon^\chi))\psi(y)\,dyd\xi \\
&= (2\pi\varepsilon)^{-d}\int_{\mathbf{R}^{2d}} e^{i\frac{(x-y)\cdot\xi}{\varepsilon}} h(x,\xi;\varepsilon,\mathrm{Op}^1_\varepsilon(\tau_\varepsilon^N))\psi(y)\,dyd\xi + O(\varepsilon^{N+1}). \qquad \text{(A.5)}
\end{aligned}$$

Now we need to compute $h(x,\xi;\varepsilon,\mathrm{Op}^1_\varepsilon(\tau_\varepsilon^N))$. Using the fact that h is polynomial in τ,

$$\begin{aligned}
h(x,\xi;\varepsilon,\mathrm{Op}^1_\varepsilon(\tau_\varepsilon^N)) &= \sum_{j,\ell=0}^{N_0} h^{j,\ell}(x,\xi)\varepsilon^\ell\big(\mathrm{Op}^1_\varepsilon(\tau_\varepsilon^N)\big)^j \\
&=: \mathrm{Op}^1_\varepsilon(\tilde{h}(x,\xi,y,\eta;\varepsilon))
\end{aligned}$$

where using compositions rules recalled in Appendix B, the symbol $\tilde{h}(x,\xi,y,\eta;\varepsilon)$ (in the (y,η) variable) can be expanded as

$$\tilde{h}(x,\xi,y,\eta;\varepsilon) = \varepsilon^\nu\sum_{k=0}^\infty \tilde{h}^k(x,\xi,y,\eta)\varepsilon^k + O(\varepsilon^\infty)$$

where the principal symbol of $\tilde{h}(x,\xi,y,\eta;\varepsilon)$ is

$$\varepsilon^\nu\tilde{h}^0(x,\xi,y,\eta) = \varepsilon^\nu\left(h^{0,\nu}(x,\xi) + h^{1,0}(x,\xi)\tau_0(y,\eta)\right). \tag{A.6}$$

We recall indeed that according to Appendix B, one has $\big(\mathrm{Op}^1_\varepsilon(\tau_\varepsilon^N)\big)^j =: \mathrm{Op}^1_\varepsilon(m_j(y,\eta;\varepsilon))$, where $m_1(y,\eta;\varepsilon) = \tau_\varepsilon^N(y,\eta)$ and

$$m_j(y,\eta;\varepsilon) = \sum_{k\ge 0}\frac{(i\varepsilon)^k}{k!}\partial_y^k m_{j-1}(y,\eta;\varepsilon)\partial_\eta^k\tau_\varepsilon^N(y,\eta) + O(\varepsilon^\infty).$$

In particular the principal symbol of $\big(\mathrm{Op}^1_\varepsilon(\tau_\varepsilon)\big)^j$ is $\varepsilon^{\nu j}\tau_0^j(y,\eta)$, which yields (A.6). We notice that due to (A.4), (A.6) implies that

$$\tilde{h}^0(x,\xi,x,\xi)=0, \tag{A.7}$$

so that $\tilde{h}(x,\xi,x,\xi;\varepsilon)=O(\varepsilon^{\nu+1})$. More generally, plugging the expansion of m_j into the formula defining $\tilde{h}$ and noticing that $h^{1,0}(x,\xi)=\partial_\tau h_{0|\tau=0}(x,\xi)$, one finds that there is $h^k(x,\xi,y,\eta)$, depending on the symbol coefficients of $h(x,\xi)$, and in the (y,η) variables on the symbols $\tau_0(y,\eta),\tau_1^-(y,\eta),\dots,\tau_{k-1}^-(y,\eta)$ only, such that

$$\forall k\geq 1,\quad \tilde{h}^k(x,\xi,y,\eta)=h^k(x,\xi,y,\eta)+\partial_\tau h_{0|\tau=0}(x,\xi)\tau_k^-(y,\eta). \tag{A.8}$$

Finally going back to (A.5) we find that

$$\mathrm{Op}^0_\varepsilon(h(\varepsilon,\lambda))\psi(x)=(2\pi\varepsilon)^{-2d}\int_{\mathbf{R}^{4d}}e^{i\frac{(x-y)\cdot\xi}{\varepsilon}}e^{i\frac{(y-y')\cdot\eta}{\varepsilon}}\tilde{h}(x,\xi,y',\eta;\varepsilon)\psi(y')\,dydy'd\xi d\eta+O(\varepsilon^N).$$

We can first perform the integration in the y variable, which creates a Dirac mass at $\xi-\eta$, and therefore we have

$$\mathrm{Op}^0_\varepsilon(h(\varepsilon,\lambda))\psi(x)=(2\pi\varepsilon)^{-d}\int_{\mathbf{R}^{2d}}e^{i\frac{(x-y')\cdot\xi}{\varepsilon}}\tilde{h}(x,\xi,y',\xi;\varepsilon)\psi(y')\,dy'd\xi+O(\varepsilon^N). \tag{A.9}$$

Now we shall construct τ_1^- ensuring that the order of $\mathrm{Op}^0_\varepsilon(h(\varepsilon,\lambda))\psi(x)$ is $O(\varepsilon^{\nu+1})$ instead of $O(\varepsilon^\nu)$. The argument will easily be adaptable by induction to $\mathrm{Op}^0_\varepsilon(h(\varepsilon,\lambda))\psi(x)=O(\varepsilon^N)$ for any $N\in\mathbf{N}$, by a convenient choice of τ_j^-, for $j\leq N-1$. We notice that in (A.9), the quantity $\tilde{h}(x,\xi,y',\xi;\varepsilon)$ can easily be replaced by $\tilde{h}(x,\xi,x,\xi;\varepsilon)$ by Taylor's formula: more precisely we write

$$\tilde{h}(x,\xi,y',\xi;\varepsilon)=\tilde{h}(x,\xi,x,\xi;\varepsilon)+(y'-x)\cdot(\nabla_{y'}\tilde{h})(x,\xi,x,\xi;\varepsilon)+O(|y'-x|^2) \tag{A.10}$$

which gives after integrations by parts

$$\begin{aligned}\mathrm{Op}^0_\varepsilon(h(\varepsilon,\lambda))\psi(x)&=\int_{\mathbf{R}^{2d}}e^{i\frac{(x-y')\cdot\xi}{\varepsilon}}\tilde{h}(x,\xi,x,\xi;\varepsilon)\psi(y')\,\frac{dy'd\xi}{(2\pi\varepsilon)^d}\\&\quad-i\varepsilon\int_{\mathbf{R}^{2d}}\nabla_\xi e^{i\frac{(x-y')\cdot\xi}{\varepsilon}}\cdot(\nabla_{y'}\tilde{h})(x,\xi,x,\xi;\varepsilon)\psi(y')\,\frac{dy'd\xi}{(2\pi\varepsilon)^d}+O(\varepsilon^{\nu+2})\\&=\int_{\mathbf{R}^{2d}}e^{i\frac{(x-y')\cdot\xi}{\varepsilon}}\tilde{h}(x,\xi,x,\xi;\varepsilon)\psi(y')\,\frac{dy'd\xi}{(2\pi\varepsilon)^d}\\&\quad+i\varepsilon\int_{\mathbf{R}^{2d}}e^{i\frac{(x-y')\cdot\xi}{\varepsilon}}\nabla_\xi\cdot(\nabla_{y'}\tilde{h})(x,\xi,x,\xi;\varepsilon)\psi(y')\,\frac{dy'd\xi}{(2\pi\varepsilon)^d}+O(\varepsilon^{\nu+2}).\end{aligned}$$

Now using (A.7), it remains to choose τ_1^- so that

$$\int_{\mathbf{R}^{2d}}e^{i\frac{(x-y')\cdot\xi}{\varepsilon}}\big(\tilde{h}(x,\xi,x,\xi;\varepsilon)+i\varepsilon\nabla_\xi\cdot(\nabla_{y'}\tilde{h})(x,\xi,x,\xi;\varepsilon)\big)\psi(y')\,\frac{dy'd\xi}{(2\pi\varepsilon)^d}=O(\varepsilon^{\nu+2}).$$

That is possible simply by looking at formula (A.8) and choosing

$$\tau_1^-(x,\xi)=-\frac{\tilde{h}^1(x,\xi,x,\xi)+i\nabla_\xi\cdot(\nabla_y\tilde{h}^0)(x,\xi,x,\xi)}{\partial_\tau h_{0|\tau=0}(x,\xi)}.$$

Note that Assumption (A.2) guarantees that τ_1^- is well defined. The argument may be pursued at the next order simply replacing (A.10) by

$$\tilde h(x,\xi,y',\xi;\varepsilon) = \tilde h(x,\xi,x,\xi;\varepsilon) \quad + \quad (y'-x)\cdot(\nabla_{y'}\tilde h)(x,\xi,x,\xi;\varepsilon) \\ + \frac{1}{2}(y'-x)\otimes(y'-x) : (\nabla^2_{y'}\tilde h)(x,\xi,x,\xi;\varepsilon) + O(|y'-x|^3)$$

and using integrations by parts again. Then the choice

$$\tau_2^-(x,\xi) = -\frac{\tilde h^2(x,\xi,x,\xi) + i\nabla_\xi\cdot(\nabla_y\tilde h^1)(x,\xi,x,\xi) - \frac{1}{2}\nabla^2_\xi : (\nabla^2_y\tilde h^0)(x,\xi,x,\xi)}{\partial_\tau h_{0|\tau=0}(x,\xi)}$$

gives that $\mathrm{Op}^0_\varepsilon(h(\varepsilon,\lambda))\psi(x) = O(\varepsilon^{\nu+3})$. We leave the rest of the induction argument to the reader. To end the proof of Theorem 3 we need to consider the case $\nu = 0$. The argument is similar to the case $\nu \neq 0$ treated above, though the formulas are slightly more complicated. We shall use the same notation as in the previous case. We start by noticing that by definition of $\tilde\tau_\varepsilon$ we have in particular for all $(x,\xi)\in\mathbf{R}^{2d}$, $\displaystyle\sum_{j=0}^{N_0} h^{j,0}(x,\xi)\tau_0^j(x,\xi) = 0$. Then as before let us write $h(x,\xi;\varepsilon,\mathrm{Op}^1_\varepsilon(\tau_\varepsilon^N)) =: \mathrm{Op}^1_\varepsilon(\tilde h(x,\xi,y,\eta;\varepsilon))$. One computes easily that the principal symbol of $\tilde h(x,\xi,y,\eta;\varepsilon)$ is (unlike the case $\nu\neq 0$)

$$\tilde h^0(x,\xi,y,\eta) = \sum_{j=0}^{N_0} h^{j,0}(x,\xi)\tau_0^j(y,\eta). \tag{A.11}$$

Note that as above one has $\tilde h^0(x,\xi,x,\xi) = 0$, so that $\tilde h(x,\xi,x,\xi;\varepsilon) = O(\varepsilon)$. One can also compute the next orders, and as in the case $\nu\neq 0$ they can be written in the following form:

$$\forall k\geq 1,\quad \tilde h^k(x,\xi,y,\eta) = h^k(x,\xi,y,\eta) + \tau_k^-(y,\eta)\sum_{j=1}^{N_0} h^{j,0}(x,\xi) j\tau_0^{j-1}(y,\eta)$$

where $h^k(x,\xi,y,\eta)$ depends on the symbol coefficients of $h(x,\xi)$, and in the (y,η) variables on $\tau_0(y,\eta),\tau_1^-(y,\eta),\dots,\tau_{k-1}^-(y,\eta)$ only. In particular we notice that

$$\forall k\geq 1,\quad \tilde h^k(x,\xi,x,\xi) = h^k(x,\xi,x,\xi) + \tau_k^-(y,\eta)\partial_\tau h_{0|\tau=\tau_0}. \tag{A.12}$$

Now that these formulas have been established, it remains to go through exactly the same computations as in the case $\nu = 0$, and we find

$$\tau_1^-(x,\xi) = -\frac{\tilde h^1(x,\xi,x,\xi) + i\nabla_\xi\cdot(\nabla_y\tilde h_0)(x,\xi,x,\xi)}{\partial_\tau h_{0|\tau=\tau_0}(x,\xi)}$$

which is well defined thanks to Assumption (A.2). The other orders are obtained exactly as in the case $\nu = 0$. This ends the proof of Theorem 3. □

Appendix B. Some well-known facts in semi-classical analysis

In this section we recollect some well-known facts in semi-classical analysis, which have been used throughout the paper. Most of the material is taken from [7], [21], [25], [30] and [29].

B.1. **Semi-classical symbols and operators.**

B.1.1. *Definitions.* We recall that an **order function** is any function $g \in C^\infty(\mathbf{R}^d; \mathbf{R}^+ \setminus \{0\})$ such that there is a constant C satisfying

$$\forall X \in \mathbf{R}^d,\ \forall \alpha \in \mathbf{N}^d, \quad |\partial^\alpha g(X)| \leq C g(X).$$

For instance $g(x,\xi) = (1+|\xi|^2)^{\frac{1}{2}} =: \langle \xi \rangle$ is an order function. Note that the variable X usually refers to a point (x,ξ) in the cotangent space $T^*\mathbf{R}^n \equiv \mathbf{R}^{2n}$, or to a point of the type (x,y,ξ) with $y \in \mathbf{R}^n$. A **semi-classical symbol** in the class $S_d(g)$ is then a function $a = a(X;\varepsilon)$ defined on $\mathbf{R}^d \times]0,\varepsilon_0]$ for some $\varepsilon_0 > 0$, which depends smoothly on X and such that for any $\alpha \in \mathbf{N}^d$, there is a constant C such that $|\partial^\alpha a(X,\varepsilon)| \leq C g(X)$ for any $(X,\varepsilon) \in \mathbf{R}^d \times]0,\varepsilon_0]$.

If $(a_j)_{j\in\mathbf{N}}$ is a family of semi-classical symbols in the class $S_d(g)$, we write that

$$a = \sum_{j=0}^{\infty} \varepsilon^j a_j + O(\varepsilon^\infty)$$

if for any $N \in \mathbf{N}$ and for any $\alpha \in \mathbf{N}^d$, there are ε_0 and C such that

$$\forall X \in \mathbf{R}^d,\ \forall \varepsilon \in]0,\varepsilon_0] \quad \Big|\partial^\alpha \Big(a(X,\varepsilon) - \sum_{j=0}^{N} \varepsilon^j a_j(X,\varepsilon) \Big)\Big| \leq C \varepsilon^N g(X).$$

Conversely for any sequence $(a_j)_{j\in\mathbf{N}}$ of symbols in $S_d(g)$, there is $a \in S_d(g)$ (unique up to $O(\varepsilon^\infty)$) such that $a = \sum_{j=0}^{\infty} \varepsilon^j a_j + O(\varepsilon^\infty)$. An **$\varepsilon$-pseudodifferential operator** is defined as follows: if a belongs to $S_{3n}(g)$, and u is in $\mathcal{D}(\mathbf{R}^n)$, then

$$\Big(\mathrm{Op}_\varepsilon(a)\Big) u(x) := \frac{1}{(2\pi\varepsilon)^n} \int e^{i(x-y)\cdot\xi/\varepsilon} a(x,y,\xi) u(y)\, dy d\xi.$$

B.1.2. *Changes of quantization.* If $a \in S_{2n}(g)$ and $t \in [0,1]$ then $a^t(x,y,\xi) := a((1-t)x+ty,\xi)$ belongs to $S_{3n}(g)$, and one defines $\mathrm{Op}^t_\varepsilon(a) := \mathrm{Op}_\varepsilon(a^t)$. When $t = 0$ this corresponds to the **classical**, or "**left**" quantization, when $t = 1/2$ this is known as the **Weyl quantization** (and is usually denoted by $\mathrm{Op}^W_\varepsilon(a) = \mathrm{Op}^{\frac{1}{2}}_\varepsilon(a)$), while when $t = 1$ one refers to the "**right**" quantization. Furthermore, if a belongs to $S_{3n}(\langle\xi\rangle^m)$ for some integer m (or more generally if a belongs to $S_{3n}(g)$ where g is a Hörmander metric [21]), then there is a unique symbol a_t belonging to $S_{2n}(\langle\xi\rangle^m)$ (resp. $a_t \in S_{3n}(g)$) such that $\mathrm{Op}^t_\varepsilon(a_t) = \mathrm{Op}_\varepsilon(a)$, and one has

$$a_t(x,\xi) = \sum_\alpha \frac{(i\varepsilon)^\alpha}{\alpha!} \partial^\alpha_\xi \partial^\alpha_\eta a(x+t\eta, x-(1-t)\eta, \xi)_{|\eta=0} + O(\varepsilon^\infty).$$

A **classical symbol** is a symbol a in $S_{2n}(\langle\xi\rangle^m)$ such that $a(x,\xi;\varepsilon) = \sum_{j=0}^{\infty} \varepsilon^j a_j(x,\xi) + O(\varepsilon^\infty)$ with a_0 not identically zero, and $a_j \in S_{2n}(\langle\xi\rangle^m)$ independent of ε. For any real number ν, the term $\varepsilon^\nu a_0$ is the **principal symbol** of the classical pseudo-differential operator $A = \varepsilon^\nu \mathrm{Op}^t_\varepsilon(a)$ (and this does not depend on the quantization). On the other hand $\varepsilon^{\nu+1} a_1$ is the **subprincipal** symbol of $A = \varepsilon^\nu \mathrm{Op}^W_\varepsilon(a)$ (in the Weyl quantization only). In the following we shall denote by $\sigma_t(A)$ the symbol of an operator $A = \mathrm{Op}^t_\varepsilon(a)$ (in other words $a = \sigma_t(A)$), and by $\sigma_P(A)$ its principal symbol.

B.1.3. *Microlocal support and ε-oscillation.* If u is an ε-dependent function in a ball of $L^2(\mathbf{R}^n)$, its **ε-frequency set** (or **microlocal support**) is the complement in $\mathbf{R}^{2n}$ of the points (x_0, ξ_0) such that there is a function $\chi_0 \in S_{2n}(1)$ equal to one at (x_0, ξ_0), satisfying

$$\|\mathrm{Op}_\varepsilon^W(\chi_0 u)\|_{L^2(\mathbf{R}^{2n})} = O(\varepsilon^\infty).$$

We say that an ε-dependent function f_ε bounded in $L^2(\mathbf{R}^n)$ is **ε-oscillatory** if for every continuous, compactly supported function φ on $\mathbf{R}^n$,

$$\limsup_{\varepsilon\to 0} \int_{|\xi|\geq R/\varepsilon} |\varphi \widehat{f_\varepsilon}(\xi)|^2 \, d\xi \to 0 \quad \text{as} \quad R \to \infty. \tag{B.1}$$

An ε-dependent function f_ε bounded in $L^2(\mathbf{R}^n)$ is said to be **compact at infinity** if

$$\limsup_{\varepsilon\to 0} \int_{x\geq R} |f_\varepsilon(x)|^2 \, dx \to 0 \quad \text{as} \quad R \to \infty. \tag{B.2}$$

B.1.4. *Adjoint and composition.* Let a be a symbol in $S_{3n}(g)$, where g is a Hörmander metric [21], and define $a^*(x, y, \xi) := \overline{a(y, x, \xi)}$. Then the operator $(\mathrm{Op}_\varepsilon(a))^* := \mathrm{Op}_\varepsilon(a^*)$ satisfies for all u, v in $\mathcal{S}(\mathbf{R}^n)$,

$$\Big((\mathrm{Op}_\varepsilon(a))^* u, v\Big)_{L^2} = \Big(u, (\mathrm{Op}_\varepsilon(a))^* v\Big)_{L^2}$$

and is therefore called the **formal adjoint** of $\mathrm{Op}_\varepsilon(a)$. In particular $\mathrm{Op}_\varepsilon^{\frac{1}{2}}(a)$ is **formally self-adjoint** if a is real. Let a and b be two symbols in $S_{2n}(g_1)$ and $S_{2n}(g_2)$ respectively, where g_j are Hörmander metrics. For all $t \in [0, 1]$, there is a unique symbol c_t in $S_{2n}(g_1 g_2)$ which allows to obtain $\mathrm{Op}_\varepsilon^t(a) \circ \mathrm{Op}_\varepsilon^t(b) = \mathrm{Op}_\varepsilon^t(c_t)$. Moreover one has

$$c_t(x, \xi; \varepsilon) = e^{i\varepsilon[\partial_u \partial_\xi - \partial_\eta \partial_v]} \left(a((1-t)x + tu, \eta) b((1-t)v + tx, \xi)\right)_{\substack{|\, u = v = x \\ \eta = \xi}} =: a \#^t b. \tag{B.3}$$

This can be also written

$$c_t(x, \xi; \varepsilon) = \sum_{k\geq 0} \frac{\varepsilon^k}{i^k k!} (\partial_\eta \partial_v - \partial_\xi \partial_u)^k \left(a((1-t)x + tu, \eta) b((1-t)v + tx, \xi)\right)_{\substack{|\, u = v = x \\ \eta = \xi}} + O(\varepsilon^\infty).$$

In particular one has $\sigma_t(A \circ B) = \sigma_t(A)\sigma_t(B) + O(\varepsilon)$. For example in the case when $t = 0$ then $\mathrm{Op}_\varepsilon(a) \circ \mathrm{Op}_\varepsilon(b) = \mathrm{Op}_\varepsilon(c)$, with

$$c(x, \xi) = a \# b = \sum_\alpha \frac{\varepsilon^\alpha}{i^\alpha \alpha!} \partial_\xi^\alpha a(x, \xi) \partial_x^\alpha b(x, \xi) + O(\varepsilon^\infty). \tag{B.4}$$

In particular if a and b are two classical symbols in the sense described above, and if one defines $A := \mathrm{Op}_\varepsilon^t(a)$ and $B = \mathrm{Op}_\varepsilon^t(b)$, then the principal symbols satisfy (if $\sigma_P(A)\sigma_P(B$ does not vanish identically) $\sigma_P(AB) = \sigma_P(A)\sigma_P(B)$.

B.2. **Semiclassical operators, Wigner transforms and propagation of energy.** One of the main interests of the semiclassical setting is that it allows a precise description of the propagation of the energy, on times of the order of $O(\varepsilon)$. We refer for instance to [15] (Section 6) for the proof of the following property (actually in the more general setting of matrix-valued operators): consider a scalar symbol $\tau_\varepsilon(x, \xi)$ defined on $\mathbf{R}^{2n}$, belonging to the class $S_{2n}(\langle\xi\rangle^\sigma)$ for some $\sigma \in \mathbf{R}$ (or more generally to $S_{2n}(g)$ where g is a Hörmander metric). We assume

moreover that $\mathrm{Op}_\varepsilon^W(\tau_\varepsilon)$ is essentially skew-self-adjoint on $L^2(\mathbf{R}^n)$. Then consider f_ε^0 an ε-oscillatory initial data in the sense of (B.1), bounded in $L^2(\mathbf{R}^n)$ and compact at infinity in the sense of (B.2), and the PDE

$$\varepsilon\partial_t f_\varepsilon + \mathrm{Op}_\varepsilon^W(\tau_\varepsilon) f_\varepsilon = 0, \quad f_{\varepsilon|t=0} = f_\varepsilon^0.$$

Then the Wigner transform $W_\varepsilon(t,x,\xi)$ of $f_\varepsilon(t)$ defined by

$$W_\varepsilon(t,x,\xi) := (2\pi)^{-n} \int_{\mathbf{R}^n} e^{iv\cdot\xi} f_\varepsilon(x - \frac{\varepsilon}{2}v)\bar{f}_\varepsilon(x + \frac{\varepsilon}{2}v)\, dv$$

converges, locally uniformly in t, to the solution W of $\partial_t W + \{\tau_0, W\} = 0$ where τ_0 is the principal symbol of τ_ε, and where the Poisson bracket is given by

$$\{\tau_0, W\} := \nabla_\xi \tau_0 \cdot \nabla_x W - \nabla_x \tau_0 \cdot \nabla_\xi W.$$

The interest of Wigner transforms lies in particular in the fact that under the assumptions made on f_ε^0, for any compact set $K \subset \mathbf{R}^n$ one has $\int_K |f_\varepsilon(t,x)|^2\, dx = W_\varepsilon(t, K \times \mathbf{R}^n)$ due to the fact that $|f_\varepsilon(t,x)|^2 = \int_{\mathbf{R}^n} W_\varepsilon(t,x,\xi)\, d\xi$.

B.3. **Coherent states.** A **coherent state** is $\Phi_{p,q}(y) := (\pi\varepsilon)^{-\frac{n}{4}} e^{i\frac{(y-q)\cdot p}{\varepsilon}}\ e^{-\frac{(y-q)^2}{2\varepsilon}}$. Any tempered distribution u defined on $\mathbf{R}^n$ may be written

$$u(y) = (2\pi\varepsilon)^{-\frac{n}{2}} \int Tu(p,q)\Phi_{p,q}(y)\, dpdq,$$

where T is the FBI (for Fourier-Bros-Iagolnitzer) transform

$$Tu(p,q) := 2^{-\frac{n}{2}}(\pi\varepsilon)^{-\frac{3n}{4}} \int e^{i\frac{(q-y)\cdot p}{\varepsilon}}\ e^{-\frac{(y-q)^2}{2\varepsilon}} u(y)\, dy.$$

This transformation maps isometrically $L^2(\mathbf{R}^n)$ to $L^2(\mathbf{R}^{2n})$. The above formula simply translates the fact that $u = T^*Tu$.

B.4. **Fourier Integral Operators.** A Fourier Integral Operator (FIO) is an operator U which can be written, for any $f \in L^2(\mathbf{R}^n)$

$$Uf(x) = \frac{1}{(2\pi\varepsilon)^{\frac{3n}{2}}} \int_{\mathbf{R}^{2n}} e^{i\Phi(x,y,\tau)/\varepsilon} a(x,y,\tau) f(y)\, dyd\tau$$

where a is a symbol of order 0, compactly supported in x and y, Φ is real valued and homogeneous of degree 1 in τ, smooth for $\tau \neq 0$. One requires also a non degeneracy condition on the phase (see [29], Chap. 9, Par. 6.11) on the support of a. Then U is continuous over $L^2(\mathbf{R}^n)$.

Acknowledgements. The authors are grateful to B. Texier for having answered many questions concerning geometric optics and semi-classical analysis. They also thank P. Gérard for interesting discussions around the study of Poincaré waves in Section 5. Finally they thank the anonymous referees for a very careful reading of the manuscript. I. Gallagher and L. Saint-Raymond are partially supported by the French Ministry of Research grant ANR-08-BLAN-0301-01.

References

[1] V.I. Arnold. *Mathematical Methods of Classical Mechanics*. Springer-Verlag, New York, 1978.

[2] D.K. Arrowsmith, C.M. Place. *An introduction to dynamical systems*, Cambridge University Press, Cambridge, 1990.

[3] G. Birkhoff & G.C. Rota. *Ordinary Differential Equations*, Ginn, Boston , 1962.

[4] J.-Y. Chemin, B. Desjardins, I. Gallagher & E. Grenier, *Basics of Mathematical Geophysics*, *Oxford University Press*, 2006, xii+250 pages.

[5] C. Cheverry & T. Paul, On some geometry of propagation in diffractive times, accepted for publication, *DCDS-A*.

[6] Y. Colin de Verdière and J. Vey, Le lemme de Morse isochore, *Topology* **18** (1979), p. 283-293.

[7] M. Dimassi & S. Sjöstrand, *Spectral Asymptotics in the Semi-Classical Limit*; Cambridge University Press, London Mathematical Society Lecture Note Series **268**, 1999.

[8] A. Dutrifoy & A. J. Majda, Fast Wave Averaging for the Equatorial Shallow Water Equations, *Comm. PDE* (2007), Vol. 32, Issue 10, pp. 1617 – 1642.

[9] A. Dutrifoy & A. J. Majda, The Dynamics of Equatorial Long Waves: A Singular Limit with Fast Variable Coefficients," Comm. Math. Sci., Vol. 4, No. 2, pp. 375 – 397, 2006.

[10] A. Dutrifoy, A. J. Majda & S. Schochet, A Simple Justification of the Singular Limit for Equatorial Shallow-Water Dynamics, in Communications on Pure and Applied Math. LXI (2008) 0002-0012.

[11] I. Gallagher, T. Paul and L. Saint-Raymond, On the propagation of oceanic waves driven by a strong macroscopic flow, `arXiv:1011.4435`, *submitted*.

[12] I. Gallagher and L. Saint-Raymond, Weak convergence results for inhomogeneous rotating fluid equations, *Journal d'Analyse Mathématique*, **99** (2006), 1-34.

[13] I. Gallagher & L. Saint-Raymond, Mathematical study of the betaplane model: equatorial waves and convergence results. Mém. Soc. Math. Fr. (N.S.). 107 (2006), v+116 pp.

[14] I. Gallagher & L. Saint-Raymond, On the influence of the Earth's rotation on geophysical flows, *Handbook of Mathematical Fluid Dynamics*, S. Friedlander and D. Serre Editors Vol 4, Chapter 5, 201-329, 2007.

[15] P. Gérard, P. Markowich, N. Mauser & F. Poupaud, Homogenization limits and Wigner transforms, *Communications on Pure and Applied Mathematics* (1997) Vol. L, 323-379.

[16] A. E. Gill, Atmosphere-Ocean Dynamics, *International Geophysics Series*, **Vol. 30**, 1982.

[17] A. E. Gill & M. S. Longuet-Higgins, Resonant interactions between planetary waves, *Proc. Roy. Soc. London*, **A 299** (1967), pages 120–140.

[18] E. Grenier: Oscillatory Perturbations of the Navier–Stokes Equations, *Journal de Mathématiques Pures et Appliquées* (1997), **76**, pages 477-498.

[19] H.P. Greenspan, *The theory of rotating fluids*, Cambridge monographs on mechanics and applied mathematics, 1969.

[20] J. Hale & H. Koçak. *Dynamics and Bifurcations*. Springer-Verlag, New-York, 1991.

[21] L. Hörmander, The Analysis of Linear Operators, Vol III., Springer-Verlag 1985.

[22] J.H. Hubbard & B.H.West. *Differential equations: a dynamical system approach II*, Texts in applied mathematics, Springer, New-York, 1995.

[23] J.-L. Joly, G. Métivier & J. Rauch, Coherent nonlinear waves and the Wiener algebra, *Ann. Inst. Fourier (Grenoble)* **44** (1994), no. 1, 167–196.

[24] A. Majda, *Introduction to PDEs and waves for the atmosphere and ocean*. Courant Lecture Notes in Mathematics, 9. New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 2003.

[25] A. Martinez, An introduction to semiclassical and microlocal analysis, Springer (2002)

[26] J. Pedlosky, Geophysical fluid dynamics, Springer (1979).

[27] J. Pedlosky, Ocean Circulation Theory, Springer (1996).

[28] S. Schochet: Fast Singular Limits of Hyperbolic PDEs, *Journal of Differential Equations (1994)*, **114**, pages 476-512.

[29] E. Stein, *Harmonic Analysis*, Princeton University Press, 1993.

[30] S. Vũ Ngoc, Systèmes intégrables semi-classiques : du local au global, *Panoramas et Synthèses* **22**, 2006.

(Ch. Cheverry) Institut Mathématique de Rennes, Campus de Beaulieu, 263 avenue du Général Leclerc CS 74205 35042 Rennes Cedex, FRANCE

E-mail address: christophe.cheverry@univ-rennes1.fr

(I. Gallagher) Institut de Mathématiques UMR 7586, Université Paris VII, 175, rue du Chevaleret, 75013 Paris, FRANCE

E-mail address: Isabelle.Gallagher@math.jussieu.fr

(T. Paul) CNRS and Centre de Mathématiques Laurent Schwartz, École Polytechnique, 91128 Palaiseau Cedex, France

E-mail address: paul@math.polytechnique.fr

(L. Saint-Raymond) Université Paris VI and DMA École Normale Supérieure, 45 rue d'Ulm, 75230 Paris Cedex 05, FRANCE

E-mail address: Laure.Saint-Raymond@ens.fr